THE UNIVERSITY OF JORDAN
AUTHORIZATION FORM

I, Jamal Nassar Omar Said, authorize the University of Jordan to supply copies of my Thesis to libraries or establishments or individuals on request, according to the University of Jordan regulations.

Signature: Jamal

Date: 12 May 2011

الجامعة الأردنيـــــة
كلية الدراسات العليا

التاريـــخ: / /

نموذج رقم (18)
إقرار والتزام بالمعايير الأخلاقية والأمانة العلمية وقوانين الجامعة الأردنية وأنظمتها وتعليماتها لطلبة الماجستير

أنا الطالب: جمال نصّار عمر سعيد الرقم الجامعي: (8070252)

تخصص: علم الحاســـوب الكليـــــة: كلية الملك عبد الله الثاني لتكنولوجيا المعلومات

عنوان الرسالة:

Retrieval of Paper-watermarks in Manuscripts Using Back-lighting and Digital Image Processing

أعلن بأنني قد التزمت بقوانين الجامعة الأردنية وأنظمتها وتعليماتها وقراراتها السارية المفعول المتعلقة باعداد رسائل الماجستير عندما قمت شخصيا" باعداد رسالتي وذلك بما ينسجم مع الأمانة العلمية وكافة المعايير الأخلاقية المتعارف عليها في كتابة الرسائل العلمية. كما أنني أعلن بأن رسالتي هذه غير منقولة أو مستلة من رسائل أو كتب أو أبحاث أو أي منشورات علمية تم نشرها أو تخزينها في أي وسيلة اعلامية، وتأسيسا" على ما تقدم فانني أتحمل المسؤولية بأنواعها كافة فيما لو تبين غير ذلك بما فيه حق مجلس العمداء في الجامعة الأردنية بالغاء قرار منحي الدرجة العلمية التي حصلت عليها وسحب شهادة التخرج مني بعد صدورها دون أن يكون لي أي حق في التظلم أو الاعتراض أو الطعن بأي صورة كانت في القرار الصادر عن مجلس العمداء بهذا الصدد.

توقيع الطالب:.. التاريخ: 2011/ 05 / 11

تعتمد كلية الدراسات العليا
هذه النسخة من الرسالـــة
التوقيع......... التاريخ.....

RETRIEVAL OF PAPER-WATERMARKS IN MANUSCRIPTS USING BACK-LIGHTING AND DIGITAL IMAGE PROCESSING

By
Jamal Nassar Omar Said

Supervisor
Dr. Hazem Hiary

This Thesis was Submitted in Partial Fulfilment of the Requirements for the Master's Degree in Computer Science

Faculty of Graduate Studies
The University of Jordan

May, 2011

COMMITTEE DECISION

This Thesis (Retrieval of Paper-watermarks in Manuscripts Using Back-lighting and Digital Image Processing) was Successfully Defended and Approved on 28/4/2011.

Examination Committee	**Signature**
Dr. Hazem Hiary, (Supervisor) Assist. Prof. of Image Processing and Document Analysis and Recognition	
Dr. Mohammad Shraideh, (Member) Assist. Prof. of Software Testing Using Evolutionary Algorithms	٢٠١١/٥/١١
Dr. Mohammad Qatawneh, (Member) Assoc. Prof. of Parallel Systems and Networks	
Dr. Ja'far Saraireh, (Member) Assist. Prof. of Mobile Network Security (Applied Science University)	

ACKNOWLEDGMENTS

In the name of Allah, the most Gracious, the most Merciful. All praise is to Allah the Almighty who has given me health, knowledge, patience, and perseverance to finish my Thesis.

My great thanks to my parents who stood all the way behind me with their support, encouragement, and prayers until this work was done. I would also like to express my special thanks to my sisters Malake and Ahlam and brother Omar for their love and support.

My deepest thanks to my supervisor Dr. Hazem Hiary for his invaluable scholarly advice, inspirations, help, and guidance that helped me through my Thesis work. I will always be indebted to him for all he has done for me, and it is indeed an honour and a pleasure to acknowledge his guidance and support in my Thesis which could not have been accomplished without him. Thank you very much for being such a fantastic supervisor.

I am indebted to Dr. Hazem Hiary for sharing some of his test images. I would also like to acknowledge the Special Collections of the University of Leeds Brotherton Library for their permission for using the manuscripts in this Thesis.

Finally, I would like to thank my friends for their generous friendship, enlightening discussion, and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

2-D	Two Dimensional
3-D	Three Dimensional
AI	Artificial Intelligence
CCD	Charge Coupled Device
CP	Control Point
dpi	Dots per inch
EVF	Estimated Verso Features
GUI	Graphical User Interface
IT	Information Technology
JPEG	Joint Photographic Experts Group
Kb	Kilo bit = 2^{10} bits = 1,024 bits
KB	Kilo byte = 2^{10} bytes = 1,024 bytes
Log-polar	Logarithmic Polar
LTR	Labelled Transmitted Recto
LRR	Labelled Reflected Recto
LRV	Labelled Reflected Verso
Mb	Mega bit = 2^{20} bits = 1,048,576 bits
MB	Mega byte = 2^{20} bytes = 1,048,576 bytes
ms	Millisecond = $1.0*10^{-3}$ of a second
µs	Microsecond = $1.0*10^{-6}$ of a second
SE	Structuring Element
MCR	MATLAB Compiler Runtime
NRMSE	Normalized Root-Mean-Square Error
R	Rotation Parameter
RGB	Red-Green-Blue
RST	Rotation-Scale- Translation
RR	Reflected Recto Image
RV	Reflected Verso Image
RRV	Registered Reflected Verso
RLRV	Registered Labelled Reflected Verso
S	Scale Parameter
SNR	Signal-to-Noise Ratio
T	Translation Parameter
TV	Television
TIFF	Tagged Image File Format
TR	Transmitted Recto Image
uint8	Unsigned 8-bit integer

LIST OF SYMBOLS

AND	Logical AND operation
Tan	Tangent trigonometric function
XOR	Logical exclusive OR operation
$\oplus$	Morphological dilation operation
$\ominus$	Morphological erosion operation
$\circ$	Morphological opening operation
$\bullet$	Morphological closing operation

LIST OF APPENDICES

RETRIEVAL OF PAPER-WATERMARKS IN MANUSCRIPTS USING BACK-LIGHTING AND DIGITAL IMAGE PROCESSING

By
Jamal Nassar Omar Said

Supervisor
Dr. Hazem Hiary

ABSTRACT

This Thesis presents a new approach that combines different digital image processing techniques. They are defined and dedicated to retrieve paper-watermarks embedded in ancient double-sided handwritten manuscripts using back-lighting reproduction technique; through removing and eliminating degradations such as foreground interference (recto writings) and background interference (verso strokes) that obstruct the shape of those valuable artifacts, acting as noise patterns that affect their legibility.

The images of the tested manuscripts are captured and digitized by a digital camera. Each paper sheet of the manuscripts is captured three times, reflected images from front and back, and a transmitted image using back-lighting scanning technique.

An area/intensity-based automatic registration method is integrated in order to align binary images of front and back sides of ancient manuscripts aimed to remove the verso strokes from the transmitted recto scans and strengthen the watermark signal. The image registration stage consists of two modules: spatial domain registration followed by log-polar registration. Among the possible image registration algorithms available, we focused on phase correlation technique and adopted the cross-correlation algorithm to estimate the parameters of geometric spatial transformations.

The resultant registration technique has been included in our overall procedure after conducting some tests that evaluated its performance on transmitted recto and reflected verso pairs in order to retrieve paper-watermarks from ancient documents within suitable or acceptable results despite of the existence of problems related to deformations caused by paper folding or crumpling.

Experimental results show that the registration technique achieves acceptable robustness to geometric distortions that include image translation, rotation, and scaling. A quality measure was considered to quantitatively evaluate the results of aligned pairs of images. On the other hand, the extracted paper-watermarks were evaluated qualitatively by judging by human eye.

Chapter 1: Introduction

"Paper-watermarks are a wonderful world of mystery and mystic."

(Grummer, A)

Ancient documents may contain watermarks. A watermark is a hidden pattern embedded intentionally in paper texture during its manufacture, which can be recognized as various shades of lightness or darkness when viewed by transmitted light; refers to thickness variations in paper (Hiary, 2008 and Biermann, 1996).

Watermarks are not always obvious on casual inspections because of the disparity of their visibility. Thence, sometimes they may require more effort in order to be extracted due to existence of some kind of interference (Biermann, 1996). Often, this interference can be presented in writings on front (recto) and back (verso) of a single paper that may overlap the watermark patterns; so, as a consequence, it makes picking them out more difficult (Hiary, 2008).

Since ancient documents are considered as valuable historical artifacts, many documents of interest are kept in private collections; so it can be difficult for watermark researchers and scholars to have access to those collections (Hiary, 2008). Various efforts have been developed in order to reproduce and exploit watermarks to assist in studying them. However, just depending on reproduction technique is usually not enough due to obstructions left on paper which interfere the watermark patterns. Thence, the need for locating and extracting their designs with minimal interference is raised (Hiary, 2008).

1.1 Research Motivation

Watermarks in papers have valuable historical information; they are very useful in the examination of paper because they were mainly used for dating, identifying paper sizes, determining paper usage and paper quality, trademarks of the paper-makers, paper mills, and locations (Hiary, 2008 and Biermann, 1996). Watermarks can also reveal facts and realities about historical connections relevant to the past by tracing and studying them. Sometimes, they can be used to correct errors in dating documents when observing twin watermarks (Hiary, 2008).

Watermarks have attracted researchers for centuries as being a human interest through identifying and classifying them (Hiary, 2008). Therefore, there is a growing need for this among librarians and antiquarians to make easier access to documental heritage by the public and scholars, to avoid damaging the original documents and that because most of them are too fragile (Tonazzini, et al., 2009).

Old manuscripts are decaying over time because of being affected by natural processes. Thence, reproduction techniques, e.g. back-lighting, that capture the complete representation and concern with the hidden information embedded in the paper in addition to the paper surface has been widely applied as one of the preservation methods by creating digital copies of them to prevent them from being lost forever (Hiary, 2008).

Back-lighting reproduction technique is characterized by being digital and can easily be available in the hands of individual scholars because it is cheaper, safer, and easier to use than other reproduction techniques as will be explained in Chapter 2 (Hiary, 2008).

Back-lighting reproduction technique requires a camera. Archiving scenes using multispectral imaging, e.g. a chromatic camera, that captures coloured scans of paper materials has an incalculable value in preserving old manuscripts for the future and enriching their digital documentation. Due to the technological advancement, the technical specifications of digital cameras are evolved over time in terms of image sampling and quantization that allows capturing higher quality images with more details. Which in turn, it becomes an important tool for analysis and documentation of old manuscripts (Tonazzini, et al., 2009).

Since back-lighting is being digital, it allows further image processing techniques to be applied easily by the means of digital computers in order to highlight watermark patterns and remove interference (Hiary, 2008). As observed, manipulating the digitized manuscripts to improve human perception or machine recognition by image processing techniques is being increasingly employed in the management of libraries and archives (Tonazzini, et al., 2009).

It is becoming increasingly clear that watermarks extraction from ancient manuscripts is being more acknowledged as a valid application of information technologies (IT) in the field of cultural patrimony.

1.2 Thesis Objectives

This research study aims:

- To develop a new approach composed of several digital image processing methods that can work successfully on challenging materials after being digitally reproduced.
- To integrate an automatic image registration algorithm that can work successfully in aligning front and back sides of ancient manuscripts.
- To attempt minimizing, as much as possible, unwanted interference that obstructs the watermark patterns caused by recto and verso writings.
- To maintain watermark features through improving the pictorial information of paper watermarks for human perspective.
- To build an easy-to-use interactive application to be, as far as possible, automatic through its underlying algorithms that reduce the amount of user intervention by performing all the steps automatically to allow different users to use it without difficulty.

1.3 Thesis Overview

The Thesis is organized as follows:

In Chapter 2, a literature overview of the most relevant concepts, terminologies and techniques pertaining to the research are presented to provide a basis for the subsequent chapters. Therefore, we first survey techniques of watermark reproduction and image registration along with their related works. In Chapter 3, we begin by presenting the source

material and digitization procedure which was followed. In Chapter 4, we describe in details the paper-watermarks retrieval approach that we have developed. In Chapter 5, a few experimental results on real manuscripts and comparison with other registration methods are shown, with discussion. Chapter 6 concludes and summarizes the contributions of the Thesis, it also suggests and lists some directions for future research. In Appendix A, enlarged samples of the ‘Mahdiyya’ copy of the Qur’an are depicted. In Appendix B, results that we have achieved are illustrated. Finally, in Appendix C, we present the user interface that we have designed to facilitate the use of our registration system.

Chapter 2: Literature Overview

We present in this Chapter a background pertaining to the research. The work presented in this Thesis is influenced by two categories of techniques: watermark reproduction and image registration.

2.1 Watermark Reproduction Techniques and Existing Related Works

A complete literature review was covered by Hiary, (2008). He gave a comprehensive presentation of the origins, state of the art, and the principle approaches used in this field.

2.1.1 Watermark Reproduction Techniques

Many digital and non-digital techniques have been developed in order to reproduce paper watermarks such as manual tracing (Heawood, E., 1950), rubbing (Haupt, W., 1881), Dylux (Allison, R. and Hart, J.), Ilkley (Schoonover, D., 1987), Phosphoresence (LIMA), Back-lighting (Edge, 2001, Moschini, 2000, Stewart, et al., 1995 and Hiary, 2008), Thermal photography (Neuheuser, et al., 2005) and Radiographic (Aken, 2003) techniques.

Tracing is a simple and cheap method, but it is time consuming, loses of accuracy, and needs skill and experience. It is subjective to amount of interference that obstructs the watermark, and depends on paper thickness. Rubbing is quick, easy and doesn't require

special equipments, but it does not produce good results, and there is a risk of damaging the original watermarked paper. Dylux has relatively low cost equipments, it is time saving and does not require dark room conditions, but it is sensitive to exposure time and distance from the visible and ultraviolet lights, unsafe due to resultant chemical gases, and so it needs for ventilated environment. It captures watermarks designs with interference and depends on paper thickness, ink opacity and light source. Ilkley is simple, quick, and the film produced can be duplicated easily, but this method requires dark room conditions and captures the watermark with other details in paper. Phosphoresence is a quick method, but it is sensitive to exposure time and distance from light, and it depends on paper thickness in addition to ink opacity. Anywise, it captures watermarks designs with interference. Thermal photographic generates good digital watermark images quickly, but it is sensitive to exposure time from thermal source and there is a risk of damaging the original watermarked paper. Radiography has the ability to display changes of paper thickness, regardless of what is printed on it, but it requires darkroom conditions and more expensive equipments, unsafe, time consuming and thus it is hard to reach for individual researchers (Hiary, 2008).

Our proposed approach is based on back-lighting reproduction technique because it differs from the other techniques in that it is purely digital; that allows to highlight watermark patterns and remove interference caused by writing ink (on both sides of paper). As a result, acquired digital images can be compared, processed, stored, retrieved and accessed easily. Also because it is simple, requires relatively low cost equipment, produces good image quality, it does not require darkroom conditions, and considered as a safe solution for capturing all details in paper. This makes it easier to preserve and store them in

digital archives to be accessed remotely (Hiary, 2008). Also, there are more web archives produced by other reproduction techniques can be found, for example, in (Koninklijke Bibliotheek and Dutch University Institute for Art History).

2.1.2 Existing Related Works of Watermark Reproduction Techniques

During the recent years, much of the previous works has been conducted on watermark extraction after being reproduced by manipulating images digitally trying to isolate clean watermark representations and to improve their appearance. Most commonly used digital image processes are mathematical morphology, edge detection, histogram enhancement, image segmentation, region extraction and image subtraction. However, the majority of these works lack the orientation of automatic parameters estimation (Hiary, 2008). The aim behind most of these works was to build watermark databases (Moschini, 2000 and the Bernstein consortium) and web archives such as (Hiary, H., Boyle R., and Ng, K.) and (Hiary, H., and Ng, K.).

Bernstein (Bernstein – The memory of papers) is an ongoing project for studying watermarks in papers. Its goal was the creation of a European integrated digital environment for paper history and expertise. Presently, it connects all European watermark databases, makes them accessible through the Internet, and provides image processing tools for measuring various aspects and characteristics of papers.

The Combination of back-lighting reproduction technique for digitization and applying digital image processing algorithms becomes an efficient method for watermark extraction. This combination has widely used in several research works.

Zamperoni (1989) suggested a database system for watermarks. The transmitted image was used in order to extract watermarks. Morphological closing operation was applied to eliminate chain lines[1] and the top-hat transform for background approximation to subtract it from the former result. He applied dilation operation for smoothing purposes, logical AND operation for image grouping, and a median filter for filtering. He managed to present the binary watermark images using contour coding, which reduces again the size of the data and makes the search and retrieval for watermarks in database even faster. However, the final output of his system suffered from interference (Hiary, 2008).

Stewart, et al. (1995) introduced two methods to extract watermarks from gray-scale images. The first one is image segmentation by using histogram thresholding. Choosing the value of a threshold depends on a ‘generate and test’ process (by making changes, and then trying again) for ink isolation from watermark. Furthermore, this technique sacrifices part of the watermark. The other technique approximates the behaviour of ink on paper. Moreover, it ignores verso writing ink and do not fully eliminate the recto writing ink. The final output of their techniques suffered from interference (Hiary, 2008).

Moschini (2000) built a watermark database. It contained watermarks with their documents information. He used some image processing techniques to enhance and highlight watermarks in the images (Hiary, 2008).

---

[1]Chain lines are a series of wide-spaced lines parallel to the shorter sides of the sheet or mould; they are used to hold laid lines during paper manufacture. These wires are thicker, and the spacing between them is larger than between the laid wires (Hiary, 2008).

Edge (2001) exchanged the camera with a flatbed scanner for the digitization procedure, where the document is placed on a glass window for scanning with a transparency adaptor to capture watermarks images in musical manuscripts. He enhanced these images in order to minimize interference using ‘Photoshop’ software (Adobe Systems Incorporated). The opacity of the reflected image is first changed and then overlay with the transmitted image. However, his approach does not work with bound manuscripts. He also relied on commercial software to manipulate the images, and a ‘generate and test’ process was used in order to pick the most suitable value for image opacity parameter at every attempt (Hiary, 2008).

Whelan, et al. (2001) used morphological operations (such as dilation, etc.), the morphological top-hat transform, and the Discrete Fourier Transform in order to remove laid lines, threshold operators (such as histogram thresholding) was also used after analyzing the histogram of the images for segmentation to remove the noise in order to extract watermarks. Furthermore, their work focused on the localization of the watermark from new paper without any foreground interference caused by writing ink and other features which may obstruct the watermark design. Their method used only the transmitted image, and did not benefit from the reflected image (Hiary, 2008).

This combination of back-lighting and image processing has also been used by Hiary (2008). His work is divided into two approaches. The first, a bottom-up approach presented a prototype to extract paper watermarks using a sequence of image processing algorithms. His approach pre-processes images to remove interferences and highlight the watermark, followed by segmentation, which achieves localization and extraction of watermark patterns and chain lines. He evaluated his approach with human opinion. The extracted

watermark designs were exported in vector form. His system gives effective results with the minimum interference compared to others' work. This approach used only the transmitted image for processing. Although, it successfully locates different kinds of watermarks in several data sets but it was limited to specific types of other data sets. The data sets are characterized by thin pen strokes, thin and uniform paper, and clear watermark designs. The results were in low interference and a strong watermark signal.

The second, a top-down approach presented a model-based technique to locating watermarks in difficult manuscripts. This approach serves as watermark image retrieval utility; it managed to remove recto material successfully, and he developed a statistical approach to locate watermark fragments from a known lexicon. Results show an excellent record of retrieval. Web archives are available on-line of the tested manuscripts as a result of his work. This approach requires both reflected and transmitted recto images for each page.

2.2 Image Registration Techniques and Existing Related Works

Image registration refers to the process of transforming the different sets of data spatially into one coordinate system. However, the majority of its methods consists of four steps: feature detection, feature matching, transform model estimation, and image re-sampling and transformation.

Image registration is an important application of image processing for comparing or integrating the data obtained from different acquisitions. The data set may consist of two or more images of the same scene taken at different times, from different sensors, or from

different viewpoints. Image Registration is actually a crucial step in all image analysis tasks in which the final information is gained from the combination of various data sources. It is used to geometrically align two images – the input image with the base image. Its objective is to bring the input image into alignment with the base image by applying a spatial transformation to the former image.

In the image registration process, the input and output images are available, but the specific transformation that produced the output image from the input generally is unknown. The objective, then, is to estimate the transformation function that compensates the geometric distortions such as shift in object positions and other factors, and then using it to register the two images (Gonzalez and Woods, 2008). As the spatial relationship between the two images is unknown and meanwhile pixel-level registration accuracy is required, the registration task for aligning two sides of a manuscript is mostly done manually (Wang and Tan, 2010).

Image registration process is based on mapping locations in one image to new locations in another image by a spatial transformation. Its key is about defining and determining the parameters of the spatial transformation needed to bring the images into alignment.

2.2.1 Image Registration Techniques

Many image registration techniques have been proposed in the literature. In general, existing image registration techniques can be categorized into two classes (Zitova and Flusser, 2003): intensity-based and feature-based methods. An extensive survey over image registration techniques can also be found in (Goshtasby, 2005).

Intensity-based method compares intensity patterns in the input and base images by means of correlation metrics or template matching techniques in order to estimate the transformation parameters – it does not require particular features to be extracted. Whole images or windows of predefined size (sub-images, templates or blocks) can be registered without attempting to detect any featured objects, where centres of corresponding sub-images can be treated as corresponding feature points.

Feature-based method finds correspondence between image features in the input and base images by establishing correspondence between a number of points. Features often could be points (particular points, region corners, or lines intersections) or group of connected points (lines, contours, and region boundaries). This method uses extracted features to estimate the registration parameters. The two sets of features in the input and base images are called Control Points (CP).

The aim of these methods is to find the pair wise (point-by-point) correspondence between the base and the input images using their spatial relations. Then, a spatial transformation is determined to map the input image to the base image.

2.2.1 Existing Related Works of Image Registration

Double-sided manuscripts are often affected by bleed-through[1] and show-through[2] effects. Therefore, many approaches have been proposed to reduce these kinds of interference to improve both human and machine readability. According to the information used, these approaches are categorized by Wang, et al. (2008) into two classes: blind segmentation and non-blind separation.

Blind segmentation refers to those approaches that attempt to clean the front side of a document without referring to the reverse side. Most of these methods treat bleed-through interference as a kind of background noise and remove them using threshold-like techniques. However, these methods are ineffective for severely degraded documents, where the intensity of bleed-through strokes is close to or higher than that of foreground texts. Under these conditions, thresholding methods either fail to remove bleed-through interference or incorrectly eliminate foreground strokes.

---

[1]Bleed-through refers to a kind of document degradation where when the ink transudes from the back side (verso) to the front side (recto) of the page. Double-sided historical manuscripts, which are preserved in many archives or libraries, are often affected by this effect (Wang and Tan, 2008 and Bianco, et al., 2009).

[2]Show-through refers to a kind of document degradation where the ink appears from the back side (verso) to the front side (recto) of the page due to the transparency of papers, may also appear in modern and well-preserved documents (Wang and Tan, 2008 and Bianco, et al., 2009).

Many other approaches addressing bleed-through correction belong to the non-blind category, which utilizes information from both sides of a document. As more information is exploited, these techniques are believed to be more accurate. However, these approaches require an accurate alignment of the recto and verso images. Since it is not easy to achieve perfect alignment, most non-blind approaches rely on manual alignment, which is slow, impractical, imprecise and involve human interactions.

Wang et al. (2009) have pointed out that perfect registration of the recto and verso images of a page is difficult for several reasons. Firstly, the fact that the registration is between foreground texts of the verso side and their partially shown bleed-through in the recto side. Secondly, different positioning of a page during image acquisition results in translational or rotational displacements between the recto image and the verso image. Thirdly, complicated local deformations such as warped or uneven surfaces could be caused by the bounding effect or due to the unevenness of the aging paper. Finally, background noise due to decolourization or stains can also affect the registration result.

Due to the mentioned difficulties, the selection of corresponding CPs for historical document registration is often performed manually by a document analysis expert. This is wearisome and time consuming when a considerably large collection of documents are to be processed. Techniques for automatic registration are therefore demanded and have been proposed. For example, area-based techniques perform registration using image patches and standard image similarity metrics such as normalized cross-correlation or sum of absolute differences, but they are susceptible to matching errors. CPs selection techniques that include corner detection and local extrema of wavelet transform detection are not applicable in historical document registration due to efficiency reasons and the fact that

corners or extreme points are often not present on bleed-through interference (Wang et al., 2009).

We present some of the works that used automatic registrations methods in order to remove bleed-through interference as follows:

Wang, et al. (2008) have proposed a two-stage hierarchical alignment technique that can efficiently and accurately align the two sides of double-sided historical documents in order to remove bleed-through interference. Their approach first coarsely aligns the two images using a pair of anchors extracted from the recto and verso images. The coarsely aligned images are then precisely aligned using block matching and Radial Basis Function (RBF) based interpolation techniques. Moreover, their method is fully automated and runs significantly faster than other reported alignment methods.

Tonazzini, et al. (2009) have proposed a system included a fully unsupervised registration method that can co-register any number of recto and verso channel maps of multispectral scans of double-sided documents. They used area-based methods, which relies on pixel intensities, such as Fourier-Mellin transform and parameter optimization.

Bianco, et al. (2009) presented a procedure aimed at improving the readability of ancient degraded documents. This procedure includes recto-verso registration, they present an effective application of the Fourier-Mellin transform for the difficult problem of registering recto-verso scans of ancient documents. The resultant registration technique has been included in their overall procedure after some tests that compare the performances of various area-based automatic registration algorithms on recto-verso images. They have implemented normalized cross-correlation in the spatial and frequency domains, the

Fourier-Mellin method, and a parameters optimization method. Although the Fourier-Mellin method has demonstrated to be the most reliable and suitable one for their application, the tests have shown its limits in presence of relevant deformations caused by folding or crumpling.

Wang, et al. (2009) presented a fully automatic method which detects specific number of corresponding CPs from historical documents for the purpose of registration. First, candidate points are located by inspecting the gradient direction maps of document images. Corresponding CPs are selected based on a dissimilarity metric that incorporates image intensity, gradient magnitude, gradient orientation and displacement. To improve the quality of the detected CPs, median filers and consistency checking are applied to correct mismatches. However, it is limited to documents contains foreground texts that are strongly written with a notable slanting writings at particular angles.

Wang and Tan (2010) presented a non-rigid registration method for restoring historical manuscripts from bleed-through distortion. Their proposed method makes use of the gradient maps of images and people's writing pattern to detect CPs. To describe the transformation between the two images of a manuscript, a mapping function consisting of a global affine and local B-splines is defined and then estimated by optimizing a cost function which takes into account image similarity and transformation smoothness.

By returning to our case, the common link between the above works and ours is that in the transmitted recto scans of our data set, the backlighting process produces an effect, as a result, in which is relatively similar to the bleed-through or show-through effect when the ink appears from the back side (verso) to the front side (recto) of the page. As though, what

distinguishes our research is of being one of the first that exploits a registration technique in order to extract watermarks.

Chapter 3: Manuscripts and Digitization

This Chapter presents a general description of the source material tested in this research and the digitization procedure used in acquiring the images of the manuscripts.

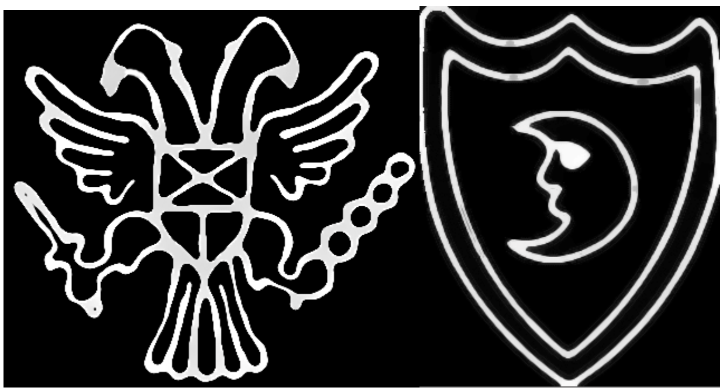
3.1 Description of the "Mahdiyya" Copy of the Qur'an

The 'Mahdiyya' copy of the Qur'an (Hiary, et al., 2008) is a digitized copy from comprehensively scanned 19th century of important and unusual Sudanese edition of the Qur'an. Each of its 346 pages contains a watermark. A part of watermark is embedded in its paper texture at the right or left hand margins in each sheet as shown in Figure 3.1. It is held by the Special Collections at the Brotherton Library of the University of Leeds (Special Collections at Leeds University Library). This data set was digitally reproduced via the back-lighting reproduction technique by Hiary (2008), as one of the contributions of his work. Our proposed approach is mainly applied and tested on this material.

(a) (b) (c)

Figure 3.1: (a) Transmitted recto scan (b) enlarged watermark area (c) enhanced illustration of the watermark signal (only for display).

The Mahdiyya' copy of the Qur'an is characterized by being the most complex data among the other digitized materials by Hiary (2008). This material is challenging for several reasons; its importance as a complete double-sided handwritten historical collection of the Qur'an, the paper sheets and writing strokes on both recto and verso sides are thick, the background is not uniform, and the watermark patterns are not clear. All these characteristics present significant interference with watermark patterns. Hence, the data is more difficult to process. Nevertheless, it doesn't suffer from bleed-through or show-through effect. The paper type used in this manuscript is laid paper. Chain lines are very hard to observe. Also, all of its papers have only one type (i.e. line/wire watermark) with two shapes (i.e. double-headed eagle watermark and moonface-within-shield countermark designs, a comprehensive knowledge of their data is illustrated in Figure 3.2) of paper watermarks in all of its pages (Hiary, 2008). The digitized images normally consist of the paper (in the centre) with a border region due to the lighting sheet during the digitization, as shown in Figure 3.3.



(a) (b)

Figure 3.2: Rough sketches of (a) a part of the design of double-headed eagle watermark and (b) the complete design of moonface-within-shield countermark.

(a) (b) (c)

Figure 3.3: Sample from the 'Mahdiyya' copy of the Qur'an, (a) reflected recto scan, (b) transmitted recto scan (c) reflected verso scan. Larger illustrations of this sample images are shown in Appendix A.

3.2 Digitization Procedure

Image acquisition is the first process of digital image processing and a fundamental step of our application. Its objective is to generate digital images from sensed data. This step is outside of our research scope because the images of the manuscripts that will be processed are already in the digital form, as a result of Hiary's (2008) work.

Hiary (2008) used an imaging system that includes FUJIFILM FinePix S1 Pro digital camera for capturing reflected and transmitted images and a green light foil for backlighting. The camera is installed against the paper to be captured at fixed distance, angle and orientation, while the light foil is replaced behind the paper, used to visualize the watermark pattern.

The images are generated by the combination of a light source and the amount of energy reflected from, or transmitted through the different elements of the paper structure (i.e. writing ink, thickness variations, watermark, etc). The sensors of the digital camera produce an electrical output proportional to light intensity while the digitizer converts these outputs to digital data (Gonzalez and Woods, 2008 and Sonka, et. al., 2008).

The used digital camera is an electromagnetic sensing device that uses sensor arrays with a 6.13 mega pixel Charge Coupled Device (CCD) imaging chip. The 'Mahdiyya' copy of the Qur'an was captured at a spatial resolution of 258 dot per inch (dpi). The quality of the digital images is usually determined by the number of samples and discrete intensity levels used in sampling and quantization by the used camera.

Each paper sheet of the manuscripts is captured three times, reflected images from front and back, and a transmitted image as shown in Figure 3.3, where the reflected recto and transmitted recto are co-registered scans. The papers are not pressed flattened on the light foil plane during the digitization procedure, e.g. via a glass window. Also the reflected recto and reflected verso were taken at different times under different conditions using the same capturing instrument.

The scans of the documents were acquired in multispectral modality. An advantage of multispectral imaging over gray-scale imaging is that each of the obtained images can further be analyzed to their image components (red, green, and blue (RGB)).

From what follows, the reflected recto, transmitted recto and reflected verso are referred to as (RR), (TR) and (RV) respectively.

In this Thesis, we assumed the following: (1) RR, TR, and RV scans of any single page of the manuscripts are available. (2) All of the scans are acquired by the same sensor or instrument. (3) The input images are of the same spatial resolution. (4) The positions of the RR and TR images are always identical to each other. (5) The front and back sides of each sheet are positioned consistently in the same way as much as possible during the digitizing stage.

Practically, the acquired images are displayed using 24-bit colour images that consist of three separate 8-bit channels, i.e. RGB channels. They are saved in an uncompressed standard graphic file format (TIFF). Thus, we expect image values to be in the range from 0 to 255. The size of each image is about 18 MB with a spatial resolution equal to 3040*2016 pixels.

Chapter 4: A Top-down Approach

4.1 Introduction

On most occasions, difficult information processing problems such as pattern extraction are approached either in a bottom-up or top-down strategy. In this research, we consider the top-down strategy to analyze and decompose the derived overview of our system by breaking it down trying to gain a new insight into its compositional sub-systems. Then, each subsystem is refined in yet greater detail, maybe into additional subsystem levels, until the entire specification is reduced to detailed basic operations (Wikipedia).

We present an approach that requires at every attempt the RR, TR and RV images of each single paper as inputs. Information resided in these inputs is utilized in order to achieve more precise output. It serves as watermarks retrieval utility that process multispectral scans of double-sided documents through applying a series of steps to extract their designs within the defined scope of this Thesis.

In our approach, inputs and outputs of the adopted image processing techniques are images which means that the type of our computerized processes is low-level processes; this involves primitive operations such as image pre-processing and contrast enhancement.

The main focus in this research is presented in adding an extension to Hiary's (2008) approach which models the effect of back-lighting in order to enhance his approach though exploiting information existed in verso scans. This has been achieved via integrating a new

stages into his model specialized in removing verso features. We have solved emerged challenges and tried to enhance the new approach in every way possible.

The overall approach operates in three main stages as shown in Figure 4.1: recto features removal, backlit-verso registration and verso features removal. Moreover, recto features removal stage can run in parallel with backlit-verso registration stage because they are independent of each other. The stage of recto features removal aims to remove the recto writing from the transmitted recto scan. Here, at this stage, the recto removal method of Hiary (2008) is used to approximate the lighting effects caused by the back-lighting acquisition between the RR and TR scans, then a differencing operation is executed.

On the other hand, the backlit-verso registration stage is used to align RV scan with its corresponding TR image in order to be passed to verso features removal stage. Its various components are detailed in Figure 4.2. Briefly, at first a pre-registration step is considered in order to enhance the input images through highlighting process, then the red component image of each input image is extracted by performing a channel selection process. Secondly, a segmentation step is took a place to separate and isolate the writings strokes from the background. Thirdly, the verso features presented in the TR image is estimated in preparation for image registration process. Finally, an image registration algorithm is used between the binary images of TR and RV of the same page in order to be transformed into one coordinate system by obtaining a registered version of the reflected verso binary image (RLRV).

Since verso features removal stage aims to remove the verso writings from TR scan, at this stage the recto removal method of Hiary (2008) is used again for the same purpose to

approximate the lighting effects caused by the back-lighting acquisition between RV and TR scans in order to be concluded by executing a differencing operation between them.

In a final step, an image grouping operation is applied between the results of the recto removal and verso removal stages in order to amplify the watermark signal which presents the final output of our overall approach. However, it must be based on accurately co-registered sets of images.

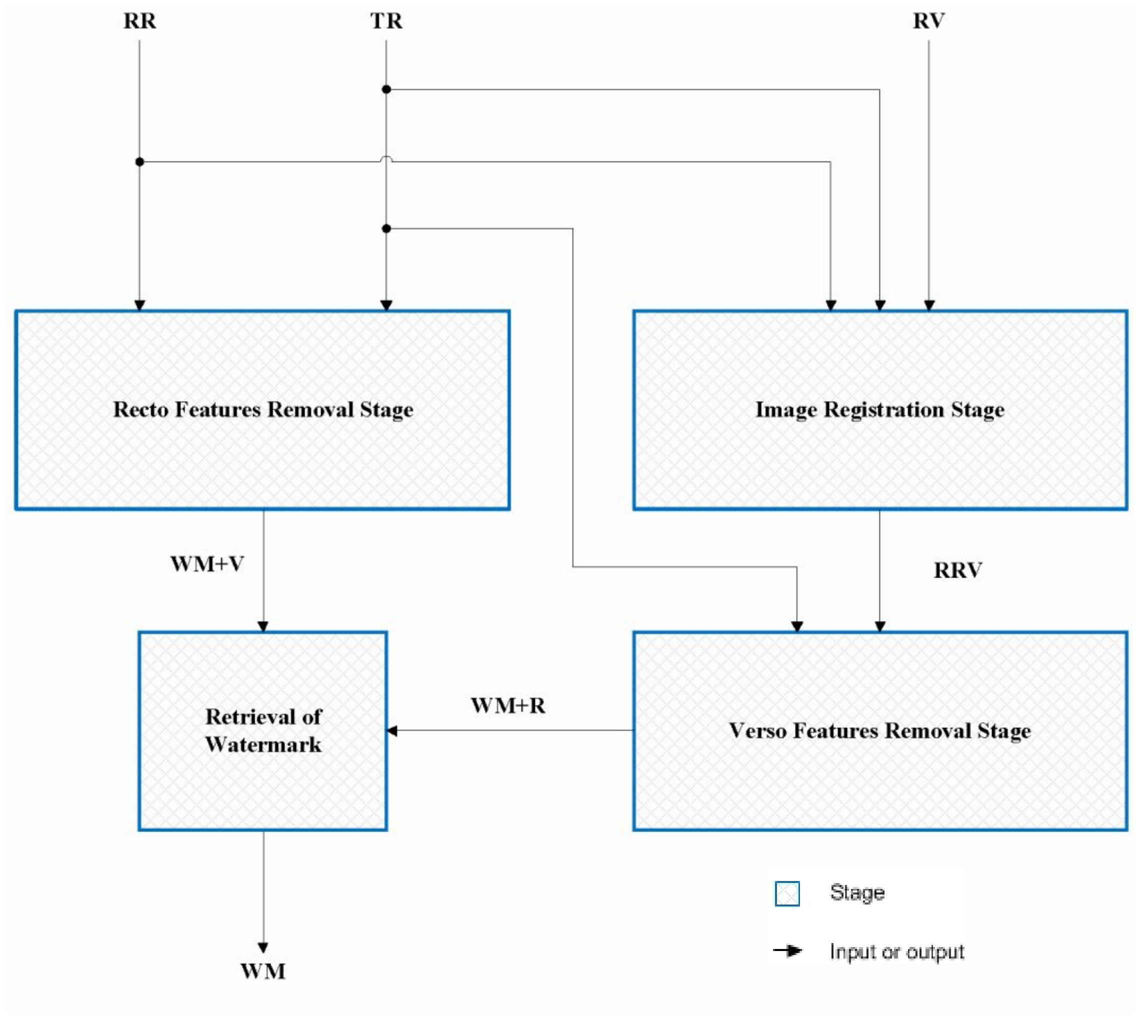


Figure 4.1: A flow chart of the paper-watermark retrieval system.

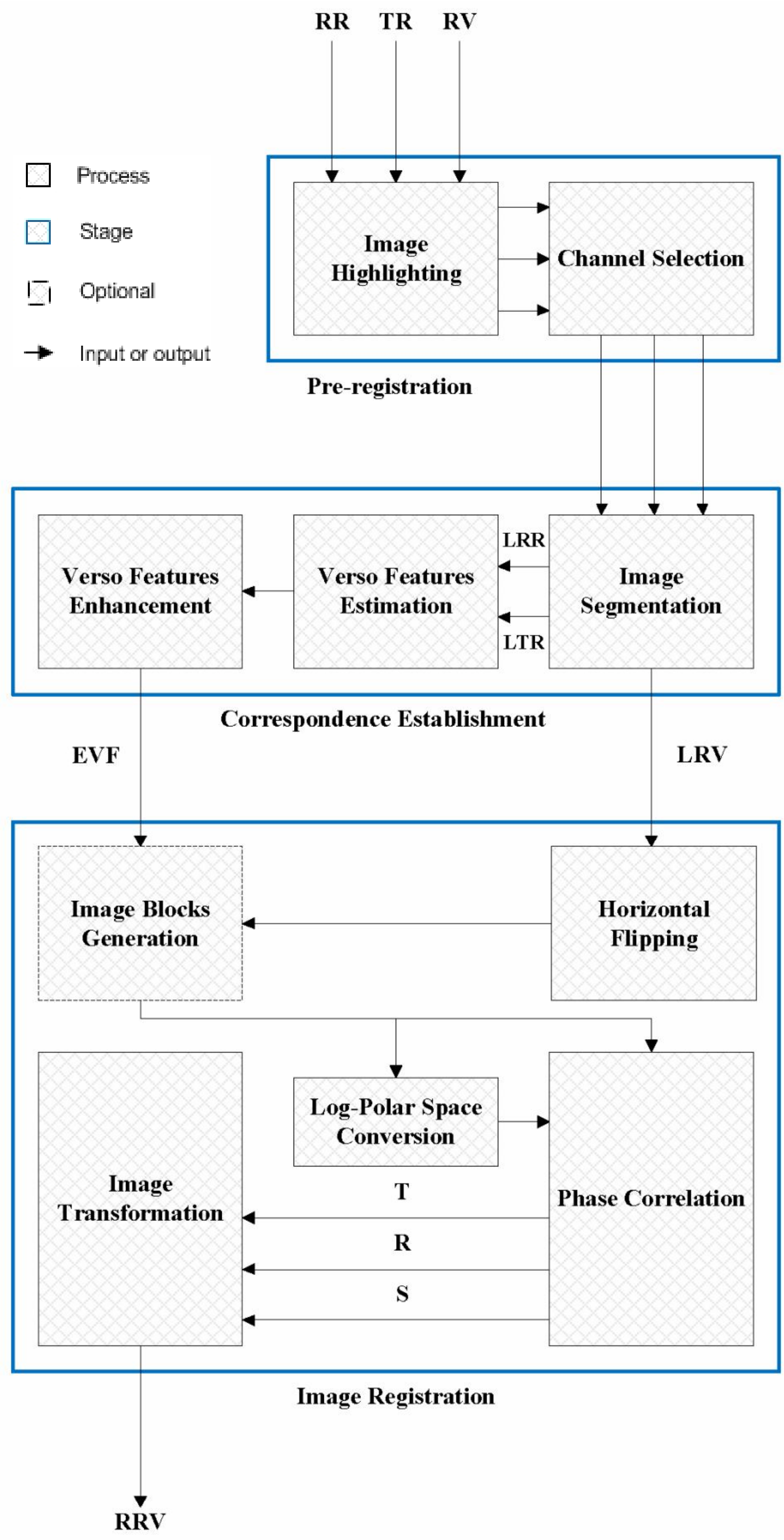


Figure 4.2: Backlit-verso registration stage.

A detailed explanation of the image processing techniques used in our approach is described in this Chapter one by one.

4.2 Recto Features Removal

Our approach starts with recto features removal stage. Hence, the recto removal section of Hiary’s (2008) model-based approach has been used in our research study to remove recto material successfully from TR images; the only difference is that we have applied it on entire images without removing the border regions of the digitized images; ensuring the simplicity of the registration procedure. His approach works well with various sets of data of different attributes and it has also been used to locate watermarks in the ‘Mahddiya’ copy of the Qur’an.

The top-down approach (Hiary, 2008) modelled a linear effect of back-lighting and exploited both of RR and co-registered TR multispectral scans of each page. The amplitude (intensity value) of the RGB channels generated at each pixel in the multispectral acquisitions is relies on the different elements of the paper structure, i.e. recto features, verso features, watermark, etc. as illustrated in Figure 4.3.

Although that RR and TR scans don’t suffer from topographical differences but they are different in colours. Therefore, Hiary (2008) hypothesized some transformation that describes the back-lighting effect. The algorithm partitions the RGB data channels of RR image into a number of clusters that contains pixels of a uniform intensity using k-means method which is controlled by some global parameter. Then, for each cluster it computes a particular transform matrix that approximates the intensity effect of back-lighting; this is

then used to subtract the transformed RGB vector at each particular pixel in RR image from the corresponding RGB in TR image in order to remove all recto information in a differencing operation which is carried out between corresponding pixel pairs as demonstrated in Figure 4.4, where the output image should contain a watermark in addition to verso features. The results of this stage are presented in Figure 4.5.

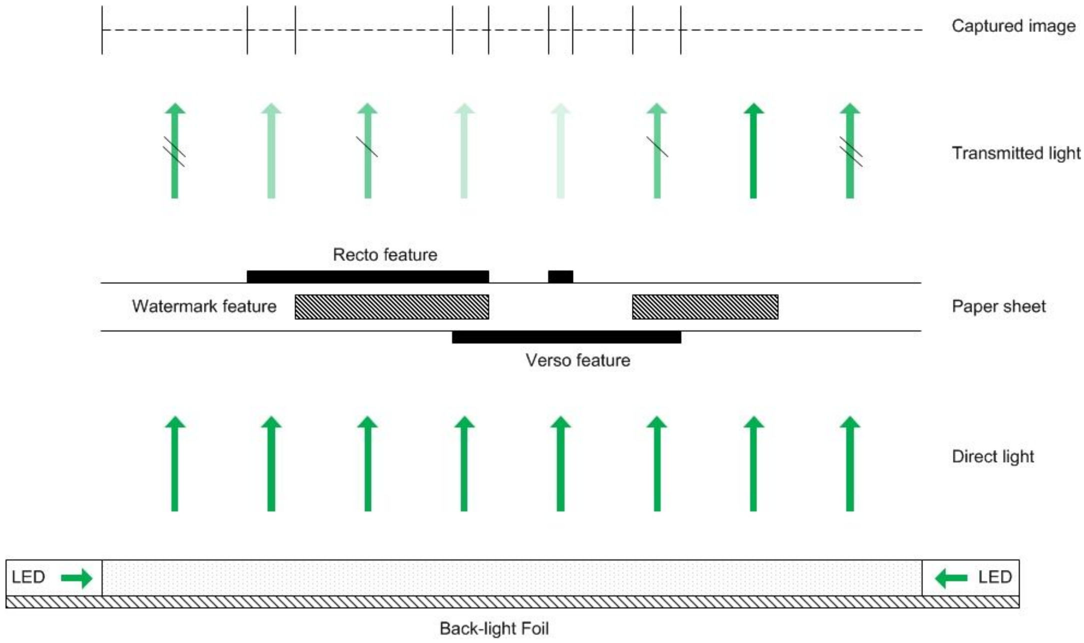


Figure 4.3: The model of back-lighting, the lowest intensity is generally caused by the combination of foreground text and interfering strokes.

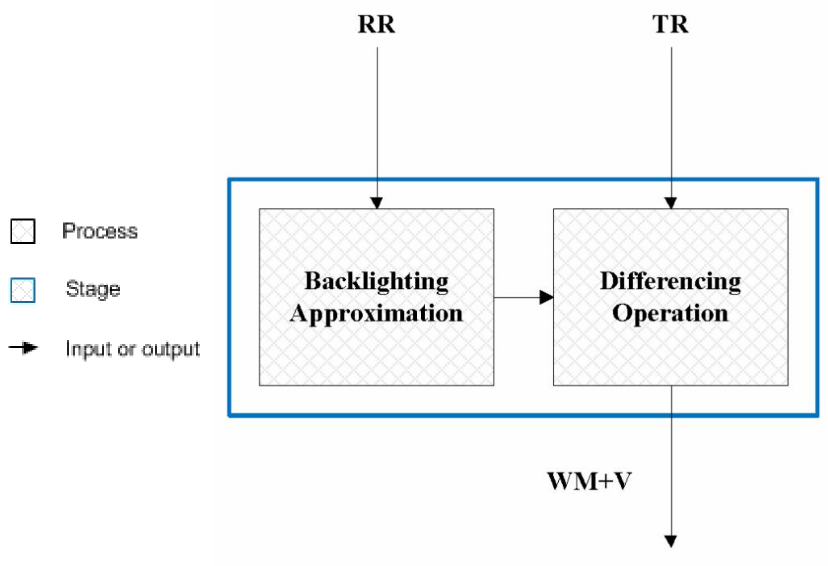
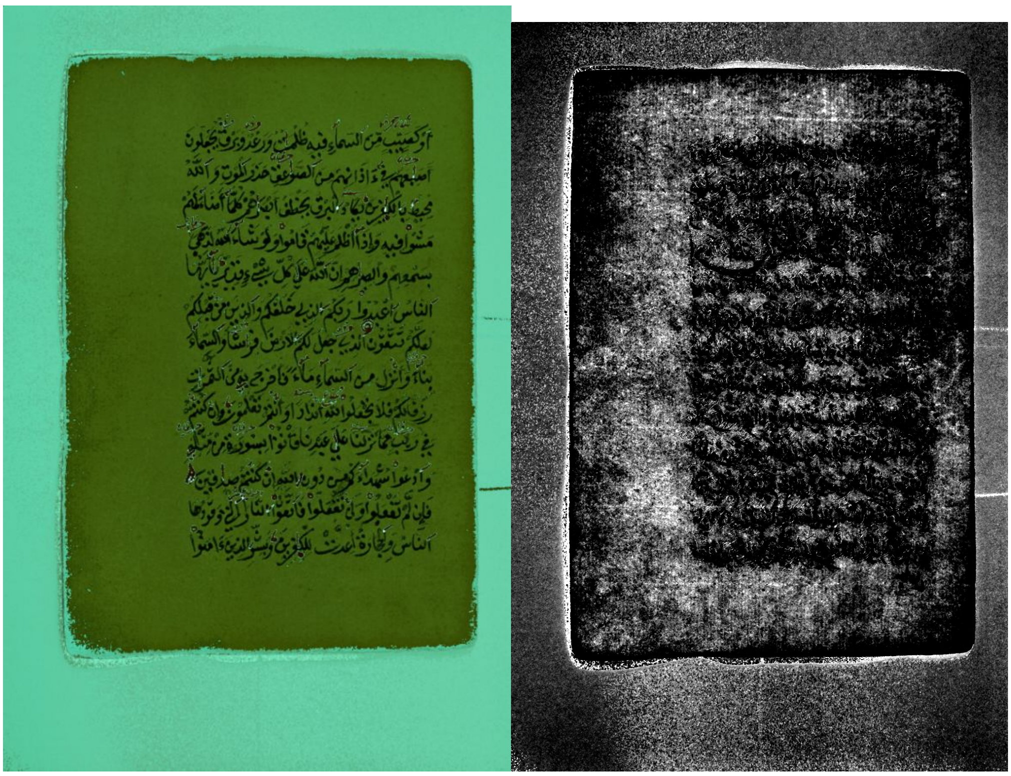


Figure 4.4: An illustration of recto removal stage.



(a) (b)

Figure 4.5: (a) Reflected recto scan after back-lighting approximation, and (b) is the result of a differencing operation between (a) and its corresponding transmitted recto image (enhanced for display). Larger illustrations are shown in Appendix B.

4.3 Backlit-verso Registration

The main steps of image registration that have been taken in order to align the two images of TR and RV are given in details in the following sections.

4.3.1 Pre-registration

An image enhancement is applied to the multispectral input acquisitions and a channel selection is then applied to the enhanced multispectral images, while the subsequent segmentation operation is typically applied to gray-scale images after just choosing a single channel. These processes act as a pre-processing stage to prepare and create binary images for the most critical process, i.e. image registration. After that, an image registration algorithm is applied to the binary images. In this case, each pair of the corresponding pixels of verso patterns in TR and RV images is aligned.

4.3.1.1 Image Highlighting

At first, image highlighting is performed to the multispectral images of TR, RR, and RV. We manipulated the input images of the scanned documents to be enhanced so that the resultant images are more suitable than the original for any further processing through our approach. However, we chose to ensure the efficacy of the segmentation algorithm by applying it on the pre-processed images as we will see in the next sections. As far as image enhancement is accomplished, the components of a document could thus be improved in a better way. Figure 4.6 shows a good enhancement result after a contrast stretching process.

(a) (b) (c)

Figure 4.6: Highlighted images of (a) RR, (b) TR and (c) RV. Larger illustrations are shown in Appendix B.

From the presented results, it is apparent that the contrast stretching process is effective as a pre-processing technique to perform the scaling necessary to spread the intensity values in the full range, especially when there is a poor contrast between writings and background to make segmentation more reliable.

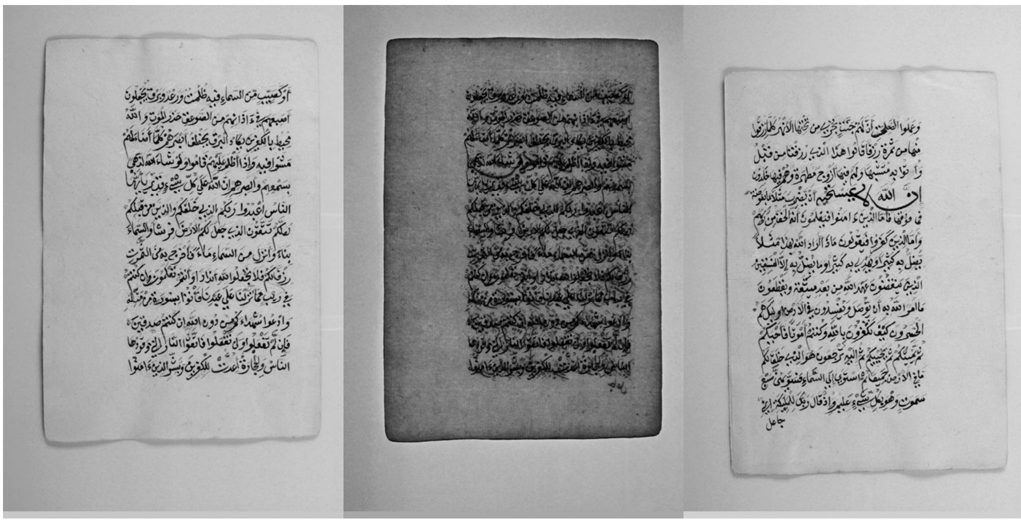
4.3.1.2 Channel Selection

TR, RR, and RV are acquired in multispectral modality. Hence, an advantage in multispectral imaging is that each of the acquisitions can further be analyzed to their three image components that present intensities of a pixel in the RGB planes that are of the same

size, allowing us to extract additional information relevant to the captured scene through exploiting the three basic component images.

Since our goal is to extract and reproduce watermark designs represented in binary – each pixel in the output images is stored as a single bit (0 or 1), we only put our emphasis on gray-scale images that are not specific for colour processing. Hence, the experiments on our proposed approach are carried out on binary images.

We decided to apply channel selection by just choosing a single component image of each multispectral image instead of RGB merge. After analysing each component image practically, we found that the red component image is the most appropriate for our application due to the good representation of the different elements of the sheet. It is worth mentioning that this process gets rid of red features from the acquisitions. The results are monochromatic images as shown in Figure 4.7.



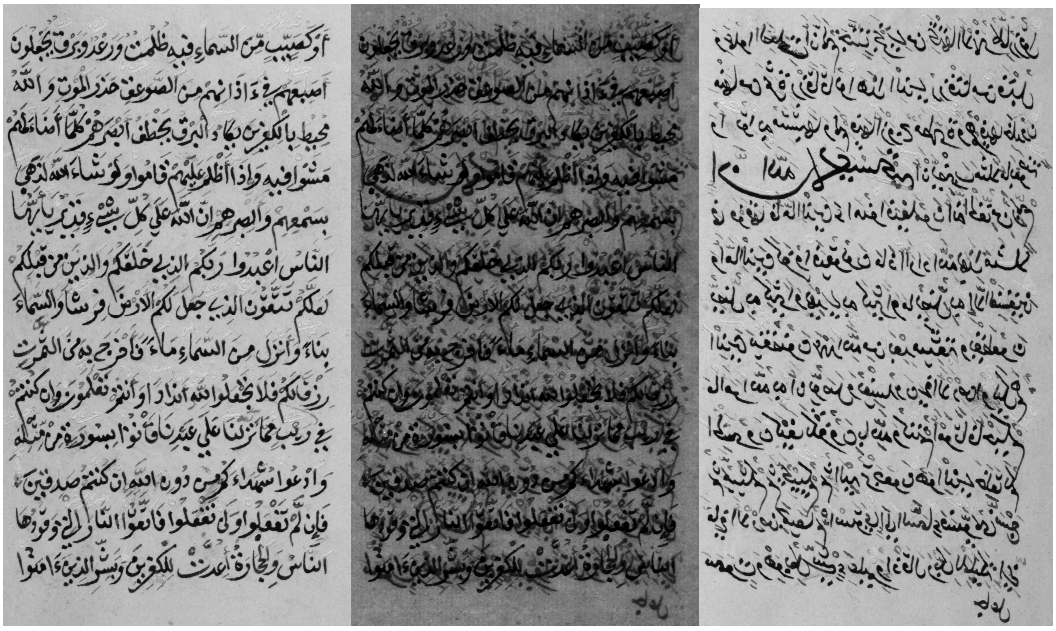
(a) (b) (c)

Figure 4.7: Monochromatic images of (a) RR, (b) TR and (c) RV after applying channel selection process. Larger illustrations are shown in Appendix B.

4.3.1.3 Image Segmentation

Image segmentation is a mid-level processing that partitions an image into regions or objects or description of those objects to reduce them to a form suitable for computer processing. Here, our case is characterized by the fact that the input is an image, but its output is attributes extracted from those images (two sets of white and black pixels).

Segmentation is typically applied to the gray-scale images to extract writing features, either recto or verso. For better estimation of dynamic global thresholds, the segmentation stage begins with region localization (cropping) of the written inscription; Figure 4.8 shows regions of interest (ROI) of RR, TR and RV.



(a) (b) (c)

Figure 4.8: ROI of (a) RR, (b) TR and (c) RV after a cropping process. Larger illustrations are shown in Appendix B.

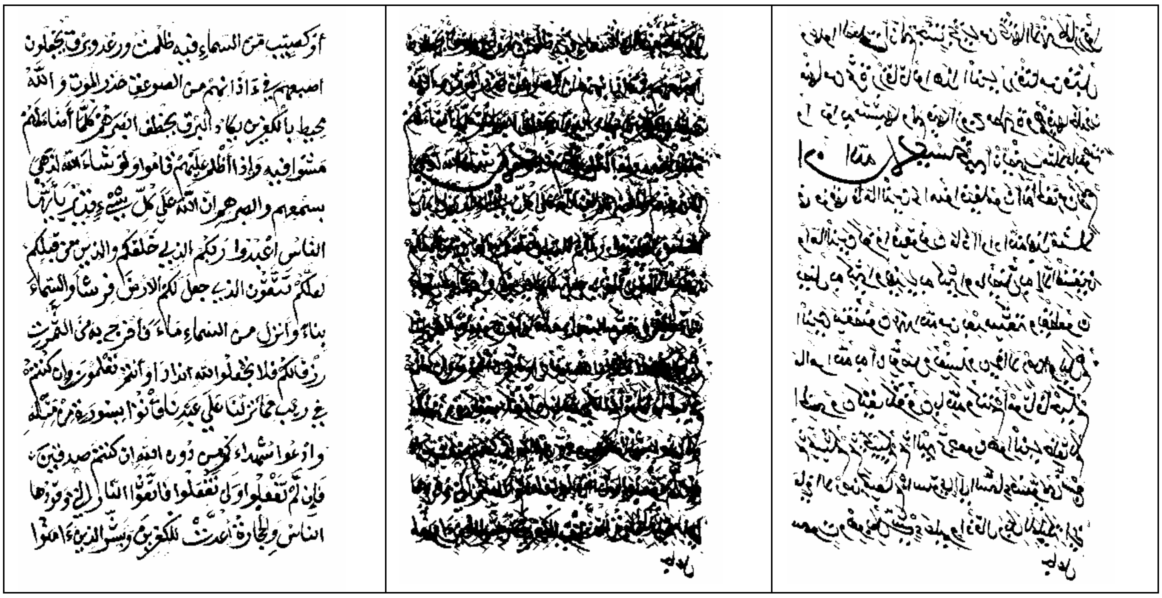
In the case of the monochromatic RR and RV images (obtained by the previous section), we aim to isolate and preserve the foreground inscription for being the most important piece of information available on the entire document that has to be useful for the registration purpose between the estimated verso strokes and the reflected verso writing as will be explained in the next sections. Here, it is obvious that the contrast between the inscription and the background is high; this leads to effective segmentation results. So, this can be obtained by just segmenting each side individually in a blind segmentation approach.

As for the monochromatic TR image, an expert eye can note the weak contrast between the recto and verso text features (the interfering strokes are relative to the corresponding original strokes of the other side of a manuscript). The intensities of verso strokes and recto text are similar (the interfering strokes have very similar intensity to the original strokes). This could make us assume that the segmentation algorithm may not be able to produce acceptable results in their separation. In this case, isolating or extracting verso writing is not trivial since their intensity is often close to the one of the main text. So we preserved both recto and verso features by just segmenting them together – this step is essential for the verso estimation process from TR images. For this purpose, the main text areas on RR and RV images are detected using a global thresholding technique.

We used a thresholding method for segmentation based on global image threasholding using Otsu's method (Otsu, 1979), since it is an efficient segmentation method to be applied to ancient documents. The effort is oriented to separate and extract text inscription from background through this segmentation method that strongly alters the appearance or aspect of the document by sacrificing and eliminating the unessential original features of the manuscript. Furthermore, the colour depth of the input images is reduced to only one bit

per pixel in order to create binary images from the monochromatic images. From what follows, the binary images obtained using Otsu's binarization method are referred to as labelled images. Figure, 4.9 depicts sample labelled images of RR, TR and RV (LRR, LTR and LRV).

Otsu's method computes a normalized intensity value, that lies in the range [0, 1], by automatically selecting it from a histogram-based thresholding and then it assumes that the individual pixels of the image to be thresholded are separated into two classes, i.e. foreground and background. Choosing of the optimum threshold value, *level*, is the key parameter in our thresholding process to minimize the intra-class variance of the black and white pixels which is expressed in terms of class probabilities and class means. Then, all pixels are replaced in the input image with luminance greater than *level* with the value 1 (white) and all other pixels are replaced with the value 0 (black).



(a) (b) (c)

Figure 4.9: Labelled images of (a) RR, (b) TR and (c) RV after segmentation process. Larger illustrations are shown in Appendix B.

In our experiments, we noticed that the thresholding algorithm is sensitive to the value of *level* used, so values of *level* for the monochromatic reflected and transmitted images are suggested from experiments to be set to 0.45 and 0.30, which result in minimum error segmentation that give the best results of our application.

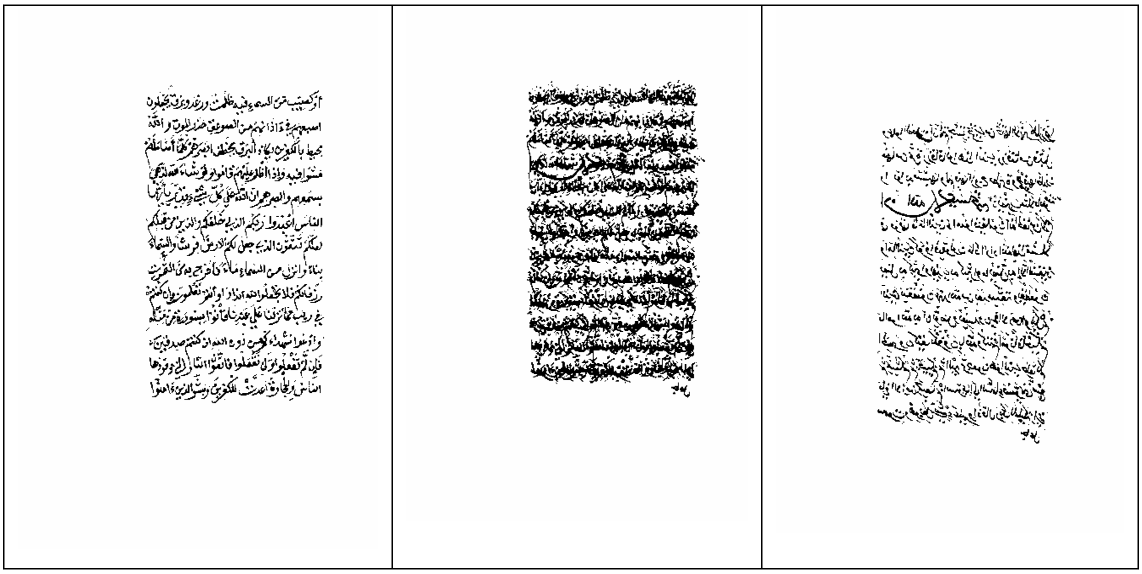
4.3.2 Correspondence Establishment

Registration of backlit-verso images is a very challenging task; since the backlit and verso sides of a page are usually different not only with respect to image intensity, but also topographically. This causes the difficulty of finding common features in the historical manuscripts. Therefore, common elements can be found for the purpose of registration among pixels of labelled estimated verso features (LEVF) and labelled reflected verso (LRV) images. As common features are difficult to be located in historical manuscripts and as pixel-level registration accuracy is required, corresponding CPs can be used for registration in this application.

Verso strokes of LTR image are estimated using a non-blind separation approach in order to find the common features between it and LRV image on which registration can be based on. The common features are often very sparse due to the subtraction operation between LTR and LRV images. Moreover, a correspondence between pixels can only be identified in the often small and sparse regions where a pattern presents in one side and also appears in the opposite side.

4.3.2.1 Verso Features Estimation

Before proceeding to the subtraction process, the area outside the cropping area in both of the paired images (LRR and LTR) is padded first as shown in Figure 4.10.



(a) (b) (c)

Figure 4.10: (a) LRR, (b) LTR and (c) LRV images after padding process. Larger illustrations are shown in Appendix B.

Each of LRR and LTR has two sets of pixels: foreground (0-valued represented in black) and background (1-valued represented in white). We applied a pixel-by-pixel logical exclusive OR (XOR) operation between LRR and LTR, which is carried out between corresponding pixel pairs. The output is the set of foreground pixels belonging to LRR or LTR, but not both, e.g. the XOR of (0,0) is 0; (0,1) is 1; (1,0) is 1; and (1,1) is 0. As noted, foreground pixels of the resultant image are composed of many small and sparse regions where a pattern present in one side also appears in the opposite side. The group of objects

of estimated verso features (EVF) are the interfering strokes which are relative to the corresponding original strokes of the back side, Figure 4.10 presents the EVF.

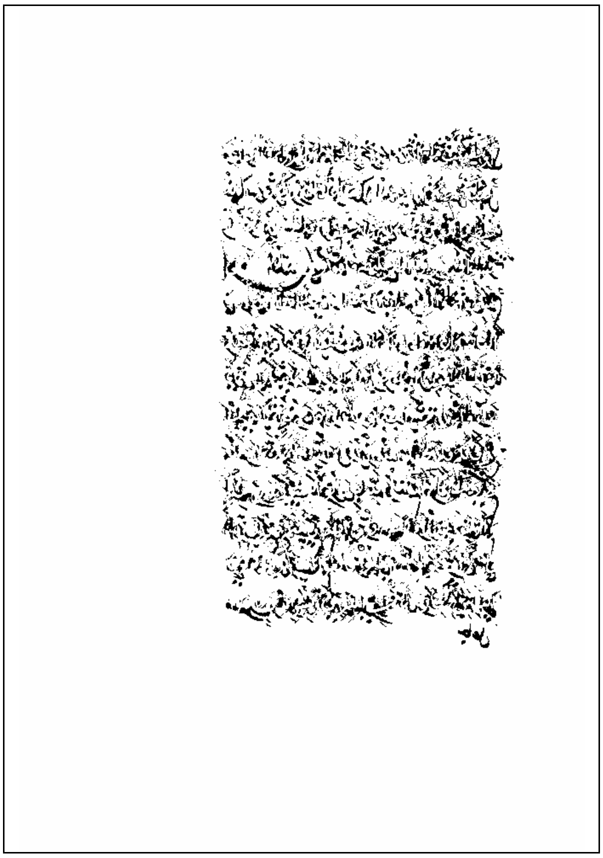


Figure 4.11: Estimated verso features (EVF) after subtraction process. Larger illustration is shown in Appendix B.

The amount of information of the EVF depends on the amount of interference or overlapping between original text and verso strokes. The more is the overlapped features, the less is the estimated verso.

4.3.2.2 Estimated Verso Enhancement

Here, we deal with Morphological operations to suppress EVF components that are not useful in its representation and description. This is achieved by applying the morphological opening because it removes the estimated verso features that are completely contained in a structuring element (SE). Opening generally smoothes the contour of an object, breaks narrow isthmuses, and eliminates thin protrusions.

The morphological opening operation is a combination of morphological erosion and dilation operation: $C = A \circ B = (A \ominus B) \oplus B$ (Gonzalez and Woods, 2008 and Sonka, et. al., 2008), where A and B are the image and the SE respectively. In practical, we used a square SE of four pixels width size in erosion and dilation operations which is enough to eliminate the pixels surrounding the recto writings in TR that have the same intensity values as the verso strokes due to horizontal ink spreading through paper tissue, this is explained in Figure 4.12. The erosion operation removes the unwanted remains of recto features and other small black regions by shrinking EV features, while the dilation operation corrects the effect of erosion on the components that are larger than the SE size by expanding EV features by the same amount. Erosion shrinks the components of an image and dilation expands them, the enhanced EVF is demonstrated in Figure 4.13.

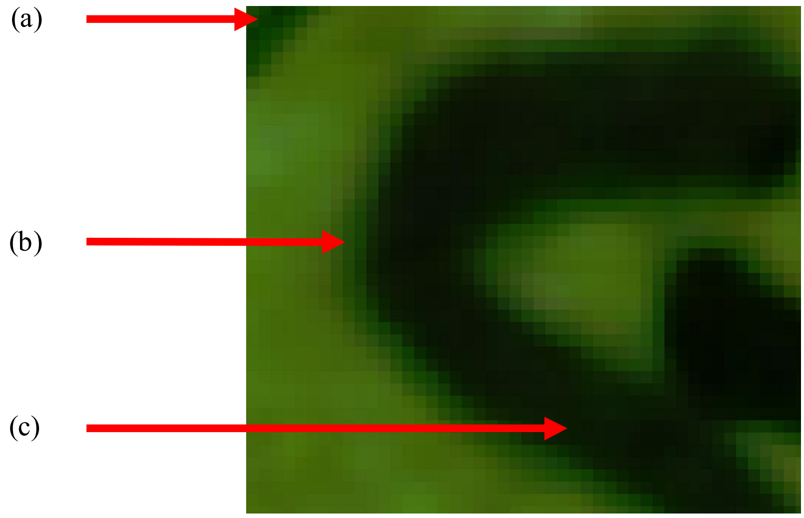


Figure 4.12: Zoomed fragment of a transmitted recto scan (a) verso stroke, (b) pixels of recto which are similar to verso pixels in intensity and (c) recto writing.

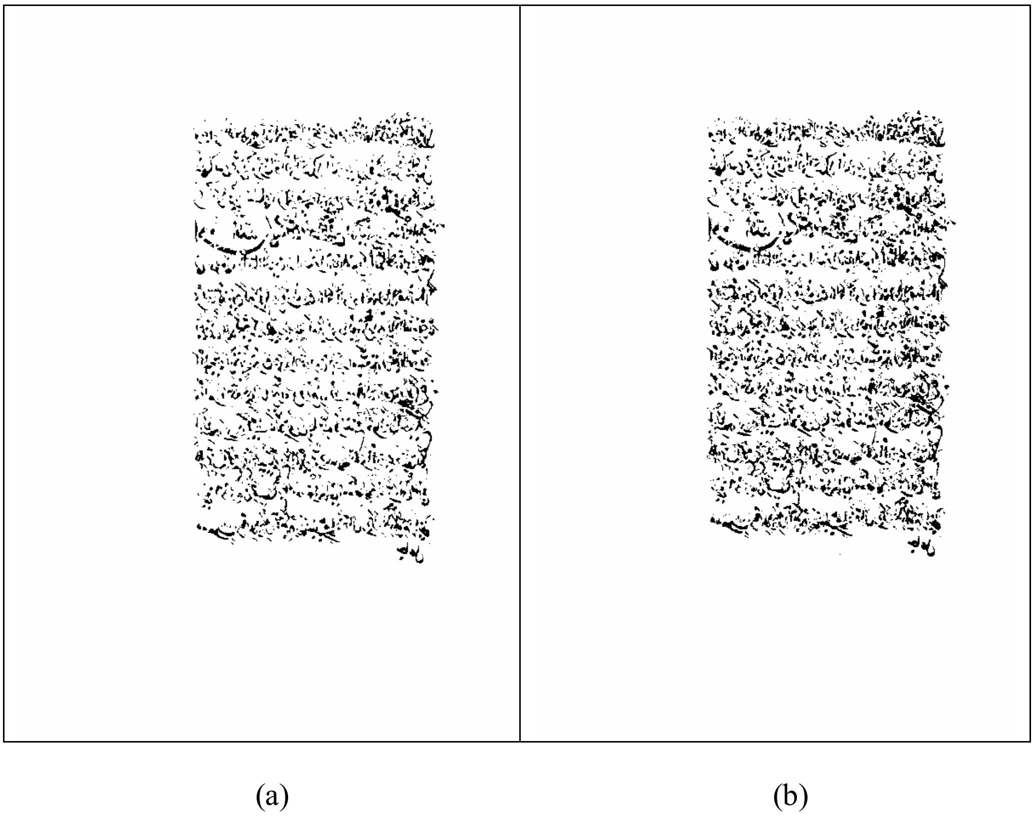


Figure 4.13: Enhanced estimated verso features after (a) erosion and (b) dilation. Larger illustrations are shown in Appendix B.

4.3.3 Image Registration

To clarify terminology, the input image is the image that we wish to transform (RV), and what we call the base image (TR) is the image against which we want to register the input. While our interest initially is on registration of binary images, LRV image is referred to as the input image and EVF image is the base image. Registration of binary images is simpler and faster than multispectral images. The process finds a correspondence between the base image and the so-called input image, and overlaps the two images through a geometric linear transformation.

In our application, the goal of image registration is to bring LRV image into alignment with EVF image by applying a geometric spatial transformation to the input image to establish geometric correspondence between them so that they are capable to be compared and analyzed. The performance of the registration process influences strongly the produced results of the subsequent steps and thus the overall output of our approach.

Our approach assumes that the front and back sides of each sheet of manuscripts are relatively positioned in the same way as much as possible during the digitization procedure. This assumption allows us to align TR and RV scans through translation transformation that serves as initial guess to the translation parameter and also speeds up the processing time of the image matching and searching processes. In general, aligning two sides of a manuscript could be difficult due to the low quality of historical manuscripts and imperfect document scanning procedure.

4.3.3.1 Geometric Distortions

In our case, we found that the differences between the input image and base image have occurred as a result of the rough alignment for the front (TR) and back (RV) sides since it is impossible to position them exactly in the same way during the digitization procedure, which causes geometrical distortions like translation and rotation displacements. We also would like to point out that our scanned historical documents were not pressed flattened on the scanning plate during the digitization procedure resulting in additional significant distortions like folding or wrapping effects occurred during the acquisition.

4.3.3.2 Images and geometric spatial transformations

The acquired images are digital, which can be defined as a two-dimensional function, $f(x, y)$, where x and y are spatial coordinates of a Cartesian coordinate system, the value or amplitude of f at any pair of coordinates (x, y) is the intensity. Whereas x, y, and the intensity values of f are all finite positive scalar quantity for each pixel determined by the source of the image, they are also processed by a digital computer.

Geometric transformations modify the spatial relationship between pixels in an image. Firstly, it transforms the spatial coordinate and assigns intensity values to the spatially transformed pixels through intensity interpolation process. The transformation of spatial coordinates is expressed as

$$(x, y) = \mathrm{T}\ \{(v, w)\}$$

where (v, w) are pixel coordinates in the input image (LRV) and (x, y) are the corresponding pixel coordinates in the transformed image (EVF).

We have used the affine transform which is a common spatial coordinate transformation; it has the following general form (Gonzalez and Woods, 2008)

$$[x\ y\ 1] = [v\ w\ 1]\ \mathbf{T} = [v\ w\ 1] \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix} \text{ .... } (4.1)$$

This transformation can rotate, scale, and translate (RST) a set of coordinate points depending on the value chosen for the elements of matrix **T**. Eq. 4.1 provides the ability to concatenate a sequence of operations as a combination of single translation, single rotation and single scale factor, all operating in the spatial coordinates of the image. Note that **T** is a geometric transformation applied to LRV image to map it from its (u, w) coordinate system to the (x, y) coordinate system of EVF. The preceding transformations relocate pixels on LRV image to new locations. To complete the process, we have to assign intensity values to those locations. This task is accomplished using intensity interpolation through considering the nearest-neighbour interpolation technique when working with these transformations.

4.3.3.3 RST Transformation

Let us consider an input image f defined over a (v, w) coordinate system, undergoes geometric distortion to produced a base image g defined over an (x, y) coordinate system.

In rotation transformation, the new coordinates of a point in the x and y spatial coordinates rotated by an angle around the z coordinate axis can be directly derived through elementary trigonometry.

Here $(x, y) = \mathbf{T}\{(v, w)\}$, where **T** is the rotation transformation applied as

$$x = v\cos\Theta - w\sin\Theta, \quad y = v\sin\Theta - w\cos\Theta$$

The transformation can be written in the following matrix form:

$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} v & w & 1 \end{bmatrix}\begin{bmatrix} \cos\Theta & \sin\Theta & 0 \\ -\sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{………………………… (4.2)}$$

In scaling transformation, if the x spatial coordinate of each point is multiplied by a positive constant Sx, then the effect of this transformation is an expansion (if $0 < Sx < 1$) or compression (if $Sx > 1$). The same can also be done along the y spatial coordinate.

Here $(x, y) = \mathbf{T}\{(v, w)\}$, where **T** is the scaling transformation applied as

$$x = c_x v, \quad y = c_y w$$

The transformation can be written in the following matrix form:

$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} v & w & 1 \end{bmatrix}\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{………………………… (4.3)}$$

In translation transformation, a translation is defined by a vector $\mathbf{T} = (dx, dy)$ and the transformation of the coordinates is given simply by

$$x = v + dx, \quad y = w + dy$$

The transformation can be written in the following matrix form:

$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} v & w & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ dx & dy & 1 \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.4)$$

4.3.3.4 Procedure Description

This Section describes in details an image registration system based on parameter estimation technique. To compute the registration parameters, we used an automatic method to extract the needed information and recover RST parameters from two images that differ in their RST. RST is a transformation expressed as a pixel mapping function that maps an input image into a base image.

We discarded the feature-based ones, since we had found that they are often insufficient to register LRV-EVF pairs, this refers to the significant differences between their features. Rather, we found good results with the area/intensity-based methods using cross-correlation. Cross-correlation is commonly used for finding translational offsets (phase shifts) between two images.

Our adopted registration method is an area/intensity-based automatic method. It spatially transforms the input (LRV) image to be aligned with the base image (EVF). It compares intensity patterns in images via similarity measures (e.g., correlation phase metric) in order to compute the values of the registration parameters. A transformation is then determined to map the input image to the base image using those parameters, thereby establishing point-

by-point correspondence between them. The goal of our method is mainly to limit the user's intervention.

The algorithm estimates the affine parameters necessary to register any two digital images misaligned due to RST. This method is based on determining the parameters of an affine transformation by finding the location of the peak that represents the maximum similarity in the correlation coefficient matrix (cross-correlation function) between corresponding pixels in EVF and LRV images. In this research work, we are only admitting affine transformations which consider and correct the geometric distortions due to small sheet displacements between successive acquisitions.

Among the possible image registration algorithms available, we focused on phase correlation technique to estimate the affine parameters. The registration algorithm works in two modules: spatial domain (Cartesian coordinate system) registration followed by Logarithmic-polar (log-polar) registration. In our implemented method, the translation parameter is computed by phase correlation in the spatial domain (Lewis, 1995), and then the rotation and scaling parameters are computed simultaneously by phase correlation in the log-polar space after converting the images to log-polar coordinates system.

The input and base images are from a single modality, i.e., binary images. Cross correlation is commonly used for registration of images in the same modality. The image registration method operates directly on the pixels of the input image in the Cartesian coordinate system in order to estimate the transition parameter. This registration method relies on pixel intensities to obtain accurate registration, and this is difficult for LRV-EVF pairs since the relevant information is often very sparse. Thus, a complete set of registration

parameters can be hard to compute because of the differences in RST. However, there is no significant relative displacements can be found between the different views of the front and back sides of the ancient documents in our acquisitions.

In particular, we used the cross-correlation in the registration method as a similarity measure to quantify the degree of similarity between intensity patterns in the base and input images which are in the same modality (Goshtasby, 2005), it do not require features to be extracted to compute the transformation parameters. They are computed iteratively using phase correlation in the Cartesian and log-polar two dimensional (2-D) coordinate systems.

The adopted metric is normalized 2-D cross correlation defined as follows (Lewis, 1995):

$$\gamma(u,v) = \frac{\Sigma_{x,y}[f(x,y) - f_{u,v}][t(x-u,y-v) - \bar{t}\,]}{\{\Sigma_{x,y}[f(x,y) - f_{u,v}]^2 . \Sigma_{x,y}[t(x-u,y-v) - \bar{t}\,]^2\}^{0.5}}$$

……………………….. (4.5)

where $\bar{f}$, $\bar{t}$ are the mean intensities of the image and template respectively, and $\bar{f}_{u,v}$ is the mean of *f*(*x*, *y*) in the region under the template. It computes the normalized cross-correlation of the image and template. The result is a matrix contains the correlation coefficients, which can range in value from -1.0 to 1.0.

A mirror image of the LRV image is acquired by flipping it horizontally with respect to the layout of the document.

The initial value of the translation parameter can be calculated in first phase translation registration to speed up the process of window searching. The coarsely aligned images are

then precisely aligned using image matching and nearest-neighbour based interpolation techniques. Using the anchor extracted from the recto and verso images as a reference feature point, we can apply the estimated values of rotation and scale parameters on it after estimating them from log-polar registration.

The algorithm of Wolberg and Zokai, (2000) is used for automatically finding the translation, the scale and rotation between the input and base images. It is based on log-polar transformations to simultaneously find the best scale and rotation parameters. Their approach is summarized below:

1. Crop central region I'_1 from I_1
2. Compute I'_{1p}, the log-polar transformation of I'_1
3. For all positions (x, y) in I_2:

    Crop region I'_2

    Compute I'_{2p}

    Cross-correlate I'_{1p} and $I'_{2p} \rightarrow (dx, dy)$

    If maximum correlation, save (x, y) and (dx, dy)
4. Scale $\leftarrow dx$
5. Rotation $\leftarrow dy$
6. Translation $\leftarrow (x, y)$

where log-polar coordinates is a coordinate system in two dimensions, and a point is identified by two numbers, one for the logarithm of the distance to a certain point (origin),

and one for an angle. A log-polar transformation maps an image I into (log r, a) coordinate space I_p. Ordinary cross-correlation in the log-polar space determines the best $d(\log r)$ and da phase shifts, which translates to scale and rotation in Cartesian space, respectively. The benefit of this new coordinate space is simple scale and rotation changes may be induced by modifying the (r, a) data).

4.3.3.5 Transform the Unregistered Image

As the final step in image registration, the original scan of RV is transformed geometrically to bring it into alignment with the TR scan. The affine transformation parameters are estimated and inferred from the registration process which was applied between their binary images in order to transform the original correspondence scan of RV. Image transform is implemented using nearest-neighbour interpolation from the original image.

4.4 Verso Features Removal

This section presents the processes used to remove verso features from TR images as illustrated in Figure 4.14. The RGB data of the registered reflected verso (RRV) which is based on a preliminary registration of the available views is transformed by back-lighting approximation. The approximated image is then used in a differencing operation with its corresponding TR. The output of the differencing operation is an image that contains the watermark and some residual noise of recto writings. Furthermore, the accuracy of verso

removal heavily relies on the registration precision. The transformed RRV scan and the result of the differencing operation are shown in Figure 4.15.

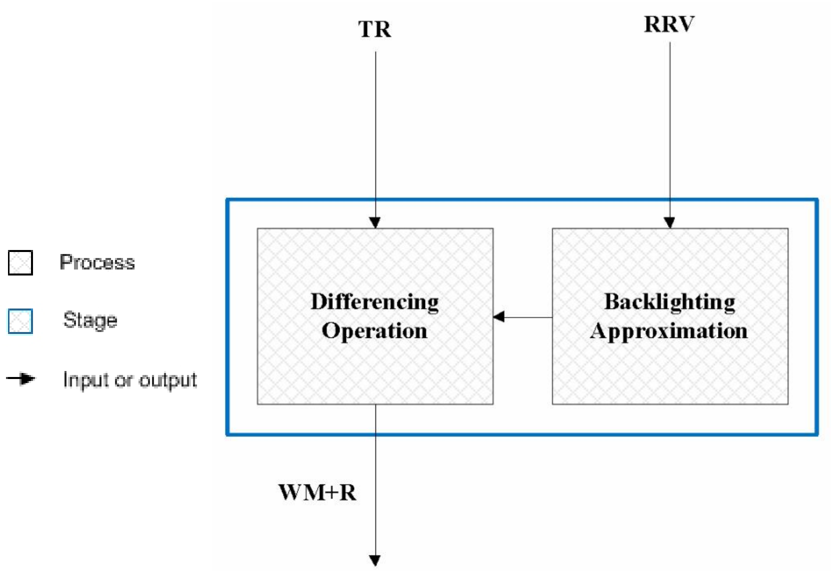


Figure 4.14: An illustration of verso removal stage.

(a) (b)

Figure 4.15: (a) Registered reflected verso scan after back-lighting approximation, and (b) is the result of a differencing operation between (a) and its corresponding transmitted recto image (enhanced for display). Larger illustrations are shown in Appendix B.

4.5 Retrieval of Watermark

A pixel-by-pixel logical AND operation is performed between the output of the recto features removal stage and the output of the verso features removal stage which is carried out between corresponding pixel pairs. The final output is the set of elements that are common between the output of the former and the latter stages, e.g. the AND of (0,0) is 1; (0,1) is 0; (1,0) is 0; and (1,1) is 1. As a result, a more visible watermark is produced. The retrieval of paper-watermark is illustrated in Figure 4.15. This can aggregate the watermark shapes and strengthen the watermark signal at the level of a single page by removing more noise and uncommon features.

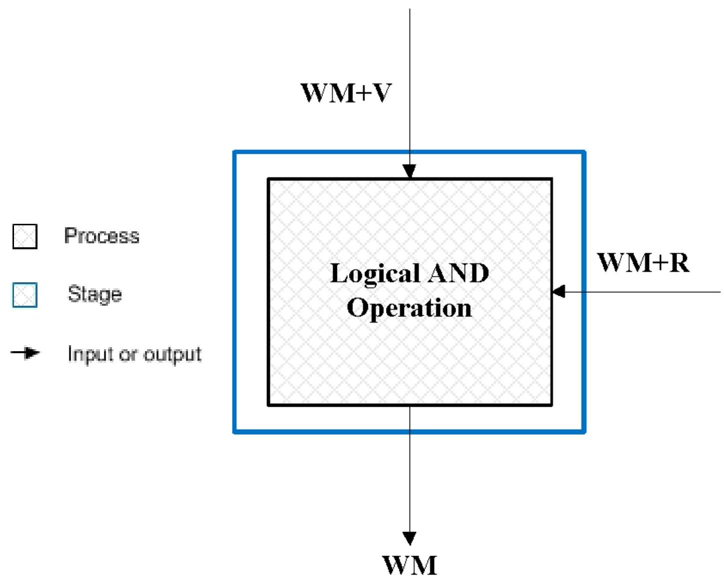


Figure 4.16: Retrieval of paper-watermark

Chapter 5: Experimental Results and Discussion

In this Chapter, the effectiveness of the registration method and the final output of the overall procedure are analyzed experimentally. The system has been implemented in MATLAB software, version 7.10.0.499 (R2010a) (The MathWorks Incorporated), by creating a wizard-style application with a Graphical User Interface (GUI) element that presents the user with a sequence of dialog boxes that lead him or her through a series of well-defined steps in order to overcome the complexity of watermark retrieving task which involves a number of operations and parameters choice – to make the steps simpler and easier to track. Figures C.1-9, in appendix C, illustrate its consecutive interfaces. The system can also be run in standalone mode that does not require for the MATLAB environment. Results were obtained using a machine with the following specifications: Intel Pentium D CPU of 3400 MHz frequency speed, 4096 MB DDR2 RAM of 400 MHz frequency speed, and under Windows Vista™ Business (service pack 2) 32-bit operating system. In practice, the registration algorithm requires around one minute and a half with image size of around 2000*1000 pixels using the previous machine. Generally, the computation time depends on the speed of the machine, the amount of memory, the content and size of the processed images. We have tested the proposed approach on 40 sheets (i.e. 120 scans) of real historical manuscripts of the 'Mahdiyya' copy of the Qur'an.

In this study, we conducted some preliminary tests on registration, the most critical step, in order to compare the performance of the implemented algorithm of the automatic method described in Wolberg and Zokai, (2000) where it was built on the basis of cross-correlation phase. To provide a basis for comparison, we concerned the way of extracting the information needed to compute the registration parameters by using the manual registration

method against the automatic method to register the double-sided documents presented in the 'Mahdiyya' copy of the Qur'an over either entire images or sub-images.

Since it is hard to achieve perfect alignment, we relied on manual alignment as the standard reference assumed to represent the maximum alignment accuracy, regardless of its negative characteristics for being slow, impractical, involves human interactions, time consuming and tedious when a considerably large collection of documents are to be processed.

The manual method is one of the principle approaches for solving the problem of image registration, it is about picking pairs of corresponding CPs in the base and input images. CPs are landmarks that can be found in both images. The minimum number of CPs required for affine transformation is at least three CP pairs in order to infer the parameters for this type of transformation. In general, the number of CPs and sophistication of the model required to solve the problem is dependent on the severity of the geometric distortion. The manual method was performed depending on our cumulative experience in manual selection of CPs during the frequent registration efforts for the pairs of historical document. It is not easy to select corresponding CPs between LRV and EVF images due to the notable differences if compared to each other (one object on one side may not appear completely on the other). For instance, Figure 5.1 demonstrates a set of pairs of corresponding CPs between the base (LRV) and input (EVF) images.

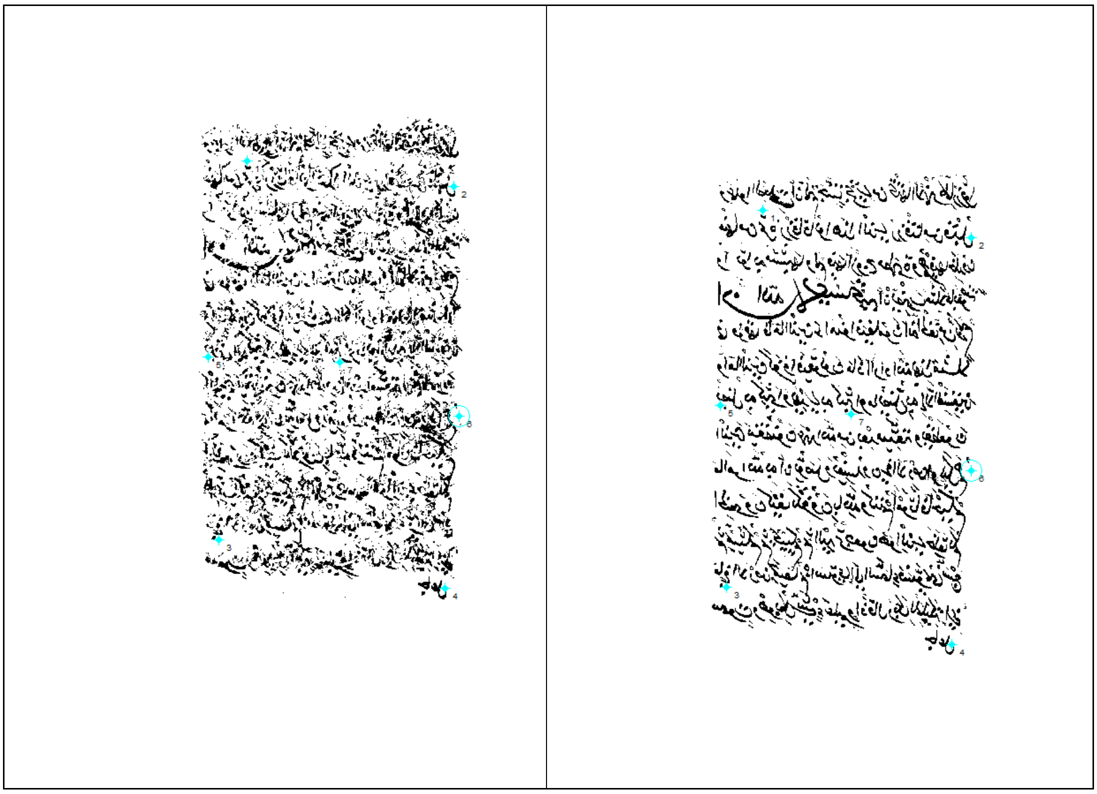


Figure 5.1: Selected pairs of corresponding CPs in the base (LEVF) and input (LVF) images.

The automatic registration algorithm can be used in two contexts: the first is to estimate the affine parameters necessary to register entire images at once consisting of the input and base images misaligned due to translation and rotation, without the need to determine the scale change. Here, alignment is performed by applying cross-correlation in Cartesian coordinate system to find the offsets *dx* and *dy* that best matches the two images, then a polar coordinate transformation is applied to the images to find the circular shift along the *a*-axis also by applying cross-correlation between the (*r*, *a*) data of the converted images. We did not take the scale parameter into account since it is useless when concerning

registration via entire images as a single block. Figures 5.2-3 present some registration results.

In the other context, images are manipulated as a group of blocks. Thus, the algorithm is used to estimate RST parameters for each corresponding blocks of the input and base images independently. It is based on log-polar transformations to simultaneously find the best scale and rotation parameters. The benefit of the division process is for seeking corresponding feature points, e.g. centres or upper left corners, between corresponding sub-images by treating them as CPs. The images are divided into equal-sizes blocks (7*4 in our experiments) automatically in order to be registered separately. Using this number of blocks in the operation helps in partial reducing of the folding (wrapping, bounding) effect which cause more errors in alignment. Then for each block on the input images, we search for its most corresponding patch to its base sub-image (block) evaluated with cross-correlation metric.

However, a common problem remains as that the accuracies of these two methods are maybe unsatisfactory due to the rigid linear transformation models they used. Figure 5.4 demonstrates this problem at the corners of the image while it shows good alignment at the centre of the image. In general, the tests have shown their limits in presence of relevant deformations caused by folding or crumpling.

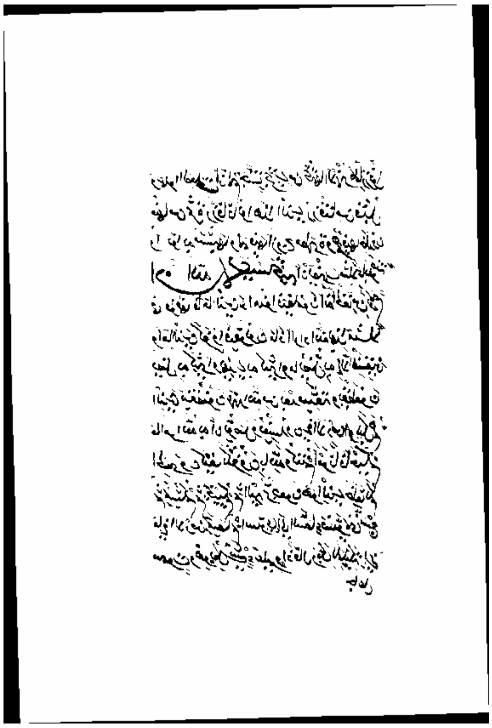


Figure 5.2: LRV after spatial transformation by rotation.

(a) (b)

Figure 5.3: (a) The RRV, and (b) the original registered reflected verso scan after spatial transformation due to translation and rotation.

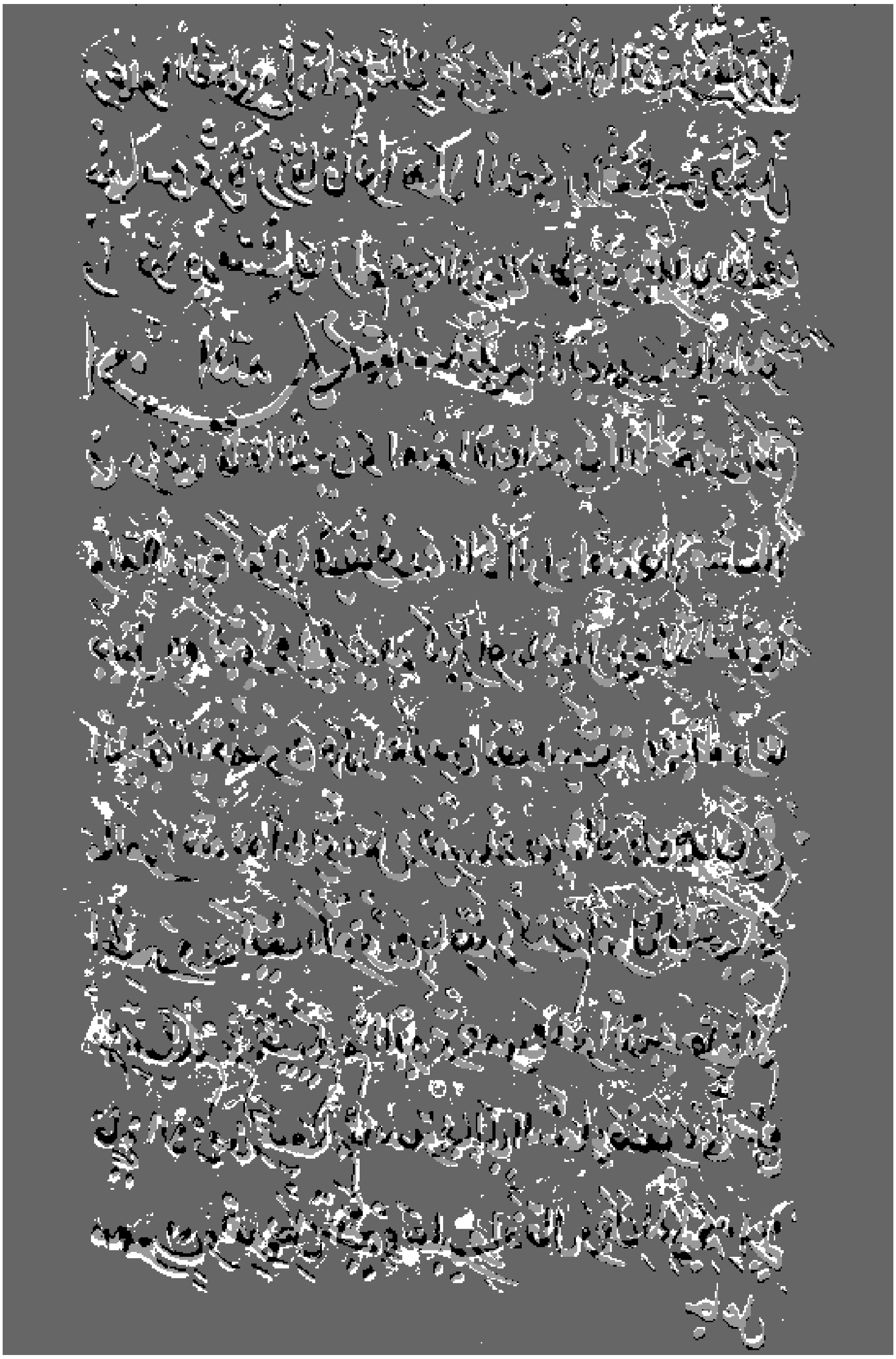


Figure 5.4: An example of an alignment; RLVF and EVF are illustrated in black and white respectively.

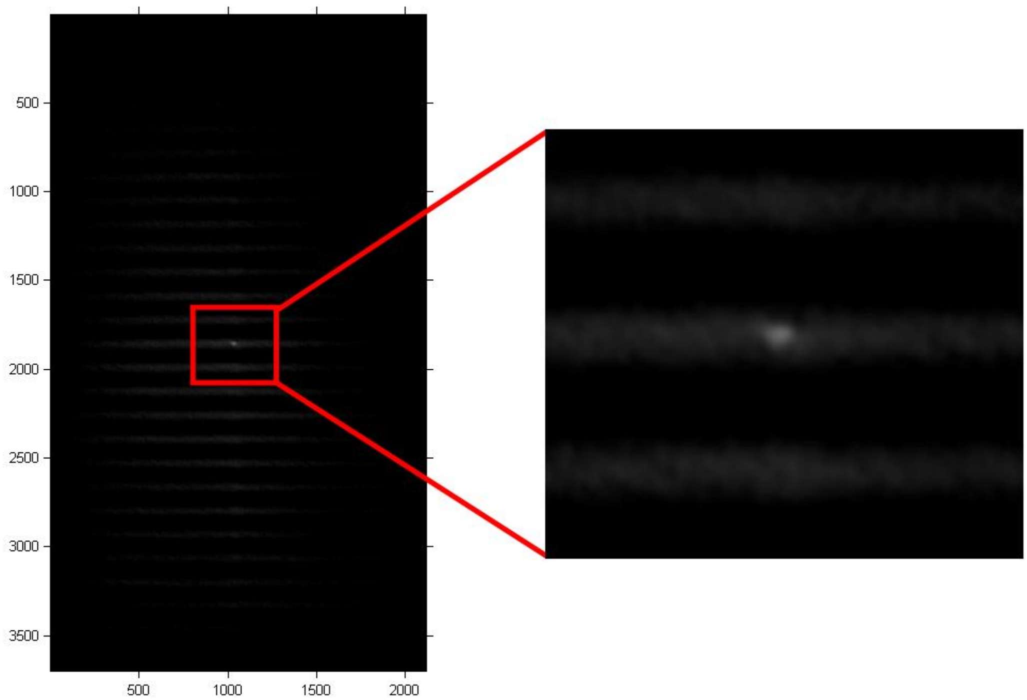
We show some examples from our extensive experimentation on real documents to arbitrate between competing methods. We also evaluated the overall effectiveness of the adopted registration algorithm on recto-verso removal and carried out a results comparison of registration algorithms of other works.

Bianco, et al. (2009) tested two methods, one computes the normalized cross-correlation in the spatial domain and the other one computes the normalized cross-correlation in the frequency but they failed in the recto-verso registration. Their dataset includes only recto and verso scans for each sheet and suffers from bleed-through effect. These algorithms, in fact, were not even able to yield a solution on recto-verso pairs, because of the significant differences between the two sides. They have also verified that in some cases none of the algorithms was able to register the recto-verso pairs. They also mentioned that this is mainly due to the notable misalignment between the two images that arises in case of a folding effect or, in general, when the page is crumpled. The failure was due to the fact that the sheets were corrugated and, probably, the affine transformation applied by the registration algorithms is not able to describe the distortions of the verso image, with respect to the recto one, induced by the sheet deformation which reflects exactly the problem of our data set.

As a quality measure for registration of a pair of images, the Normalized Root-Mean-Square Error (NRMSE) is usually adopted in the related works to evaluate quantitatively the results of registration, defined as follows (Guizar-Sicairos, Bianco, et al., 2008):

$$E^2 = 1 - \frac{\max_{u,v} \left| r_{fg}(u,v) \right|^2}{\sum_{x,y} \left| f(x,y) \right|^2 \cdot \sum_{x,y} \left| g(x,y) \right|^2} \quad \text{..........................} \ (5.1)$$

where $f(x, y)$ and $g(x, y)$ are the registered backlit and verso images respectively, and $r_{fg}(u, v)$ is their correlation coefficient (cross-correlation function). However, it is noted that the algorithm works successfully and exhibits very different behaviour than in Bianco, et al. (2009) and this is due to the availability of clear independent recto and verso writings in our dataset. Thus, evaluation of the above equation requires solving the more general problem of sub-pixel image registration by locating the peak of the cross-correlation $r_{fg}(u, v)$. For instance, it is illustrated in Figures 5.5-4.



(a) (b)

Figure 5.5: (a) A visual 2-D presentation of the resulting matrix that contains the correlation coefficients, (b) the position of the significant peak of the maximum correlation coefficient (zoomed).

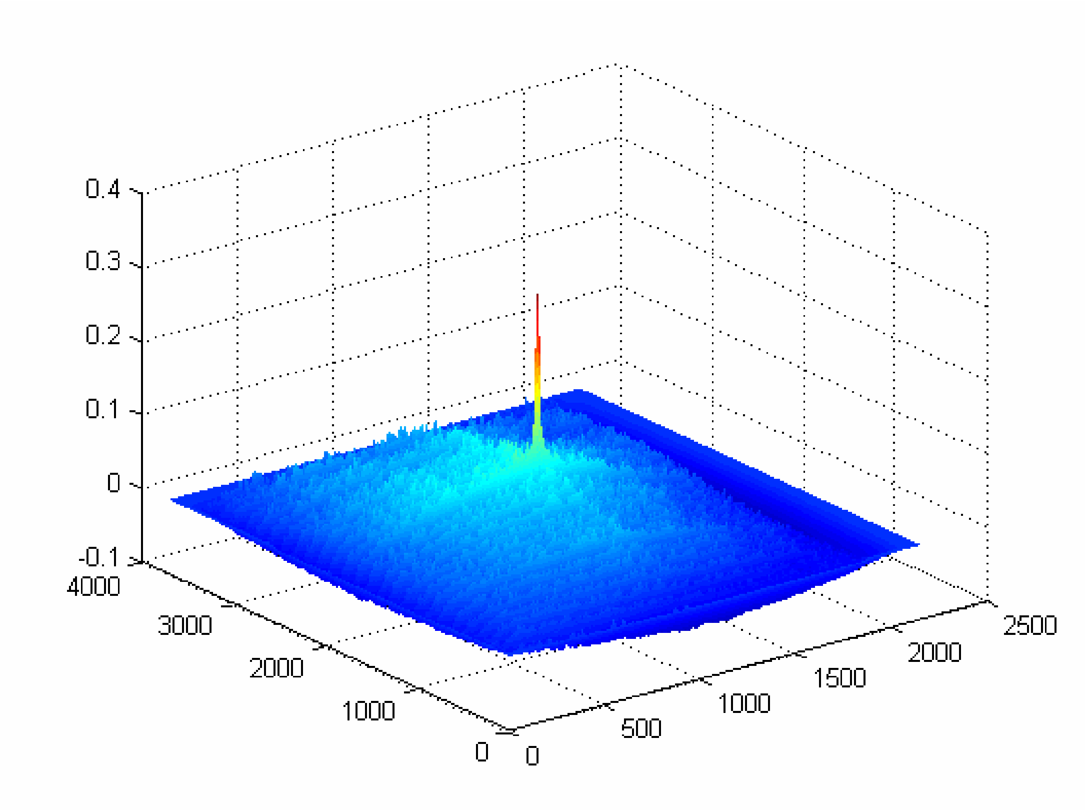


Figure 5.6: A three-dimensional (3-D) shaded surface plot of the resulting matrix that contains the correlation coefficients.

Generally, the root-mean-square error is commonly selected measure regarding the quantitative evaluation in the results of image registration. The correlation coefficient and root-mean-square error are interrelated. The correlation coefficient describes the strength of the linear relationship between registered and original images. The correlation is computed over a spatial domain for both of the images. It is a continuous parameter – that is, it is sensitive to the finest details of each registered versus original case. If the registered (f) and original (o) are standardized and denoted as z_f and z_o, the coefficient of correlation between f and o, r_{fo}, is defined as (Barnston, 1992)

$$r_{fo} = \frac{\sum_{i=1}^{N}(z_{fi}z_{oi})}{N} \quad \text{.......................... (5.2)}$$

where N is the number of space elements. The i denotes the element number. The correlation would not be different if computed without first standardizing f and o; however, the complete correlation formula would be required in which standardization is accomplished using the means and standard deviations of f and o. If there were an exact linear functional relationship between registered and original images (implying perfect registration), the correlation would be at its maximum possible value of 1.0.; if there were no linear predictability whatsoever, it would be zero.

Root-mean-square error (RMSE) is also a verification measure, which is defined as the square root of the mean of the squared differences between corresponding elements of the registered and original images (Barnston, 1992):

$$RMSE = \left[\frac{\sum_{i=1}^{N}(f(x,y) - g(x,y))^2}{N}\right]^{\frac{1}{2}} \quad \text{.......................... (5.3)}$$

where $f(x, y)$ and $g(x, y)$ are RLRV and EVF images respectively, and N is the number of space elements. In our work, we have used Eq. 5.3 as the quality measure for registration to evaluate quantitatively the results of registration between the binary images of EVF and RLRV. In addition, Eq. 5.2 can be used to produce equivalent results of correlation coefficients as Eq. 4.5 for our binary images.

There can be an exact one-to-one relationship between the correlation coefficient and the RMSE parameter when the latter is formed using standardized registered and original

images. Our experiments show that the raw numbers of correlation coefficients and RMSE are affected by the amount of the estimated verso features and the strength of the folding effect. We considered the values of correlation coefficients and RMSE that we had obtained as arbitrary numbers, which are not necessarily to be meaningful to us. Therefore, we relies on the subtraction process output between the corresponding values of correlation coefficient or RMSE for the manual method and the automatic method in order to judge on the performance and accuracy, based on the two metrics, of the automatic registration method.

With respect to paper-watermark retrieval experimentations, we employed the automatic method in backlit-verso registration in the overall procedure in order to test the final result. To evaluate the proposed method qualitatively, visual assessment can be performed by judging by an inspector's eye to decide whether an extracted watermark pattern is 'acceptable' or not. For most tested images, the proposed approach shows encouraging results. This can be presented in Figure 5.7.

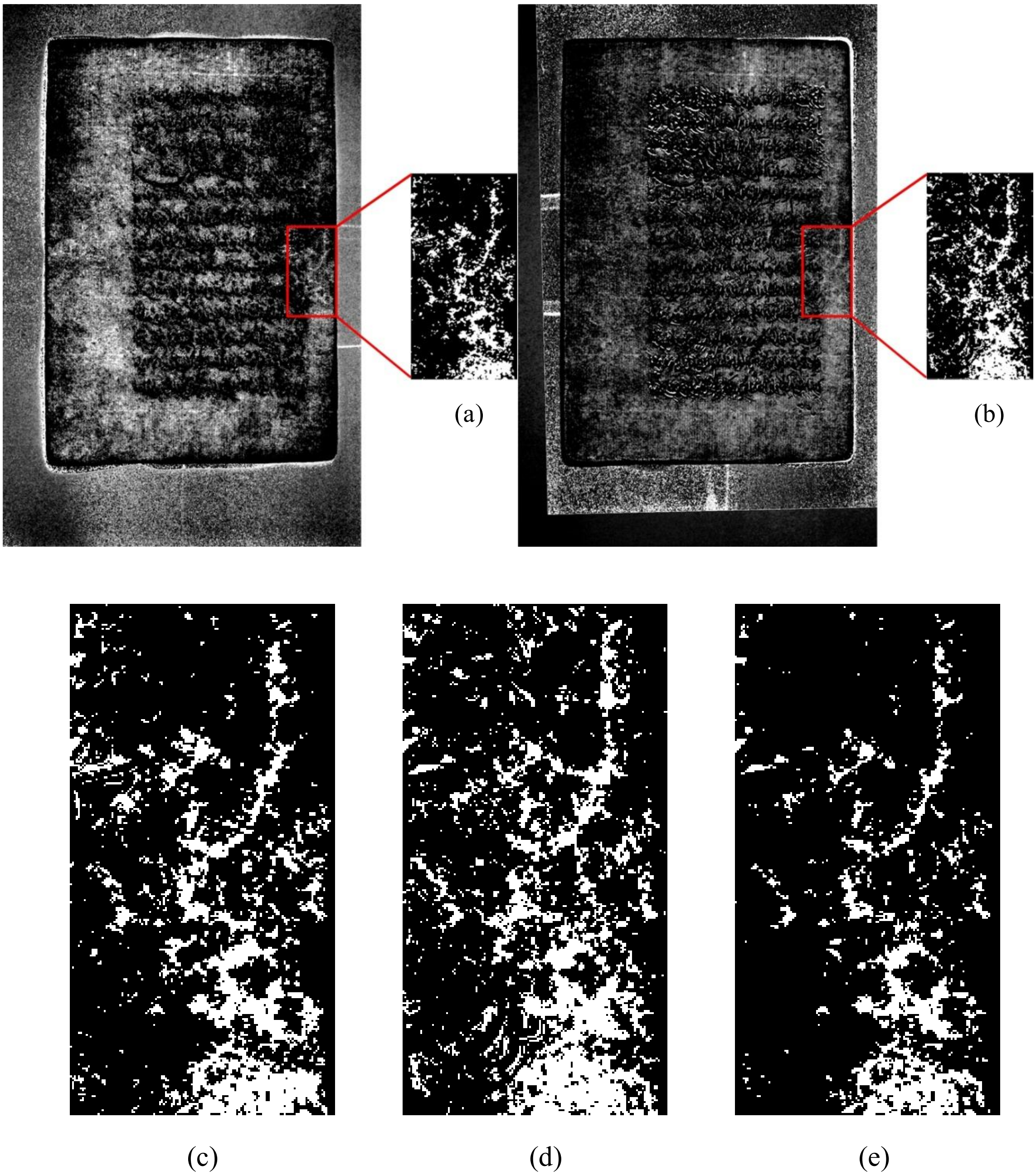


Figure 5.7: (a) is a binary sample output of the recto removal stage, (b) a binary sample output of verso removal stage, (c) zoomed illustration of (a), (d) zoomed illustration of (b), and (e) the final output of the paper-watermark retrieval system after performing the logical AND operation between (c) and (d), it shows a more visible watermark design by removing more noise and uncommon features in order to strengthen the watermark signal and aggregate the watermark shape. Complete illustrations of this sample images are shown in Appendix B.

Chapter 6: Conclusions and Recommendations

6.1 Conclusions

In our approach, image registration may fail unless the two images are relatively misaligned by a slight difference in rotation and translation. It is based on binary pairs which reduces the size of the data and makes it even faster, unlike the case of multispectral pairs, which many images need to be aligned. This can become more difficult to handle and have a tendency to subjective errors.

During our, research study, we have discovered that some pages of 'Mahdiyya' copy of the Qur'an suffers from distortion such as paper folding that makes dissimilar misalignment transformation from one region to another along each individual paper sheets. The experimentation results have shown the limitation of the registration method in presence of relevant deformations caused by folding or crumpling.

We also found that perfect registration of TR and RV images of a sheet is however difficult for several reasons. Firstly, image registration is actually between foreground texts of the verso side and their estimated verso features in the TR side. Secondly, different positioning of a page during image acquisition results in translational or rotational displacements between TR and RV images. Finally, complicated local deformations such as warped or uneven surfaces caused by the bounding effect increase the complexity of alignment. As though, what distinguishes this study is of being one of the first that exploits a registration technique in order to extract watermarks.

6.2 Recommendations

Experiments on real historical document images have shown encouraging results and further improvements can be made by exploiting more sophisticated registration methods. The alignment procedure can be improved using other registration methods to handle other types of local and global geometric transformations along the complicated local deformation, especially for documents which are not flat due to age, handling, and physical construction. The used model of back-lighting relies on linear relationship, so other better models can possibly be proposed through utilizing quadratic or cubic approximations. It can also be considered adding a post-processing routine to represent the extracted watermark after being retrieved so that the quality of resultant their images could be further improved.

BIBLIOGRAPHY

Adobe Systems Incorporated, Adobe Photoshop CS5 software. Retrieved on April 27, 2011, from **http://www.adobe.com/products/photoshop/photoshop/**.

Aken, J., (2003), **An improvement in Grenz radiography of paper to record watermarks, chain and laid lines**, Studies in conservation, 48(2), 103-110.

Allison, R. and Hart, J., Paper watermarks, paper making, and dylux proofing paper. Retrieved on April 27, 2011, from **http://www.watermarkarchive.org/watermarkinitiative/**.

The Bernstein, The memory of paper, Austria. Retrieved on April 27, 2011, from **http://www.bernstein.oeaw.ac.at/**.

Bianco, G., Bruno, F., Tonazzini, A., Salerno, E., and Cosole, E., (2009), **Recto-verso registration, enhancement and segmentation of ancient documents**, IEEE 15th International Conference on Virtual Systems and Multimedia.

Biermann, C., (1996), **Handbook of Pulping and Papermaking**, (2nd ed.), San Diego, California, USA: Academic Press, 171.

Barnston, A., (1992), **Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score**, Weather and Forecasting, 7, 699-709.

Dutch University Institute for Art History, International Database of Watermarks and Paper Used for Prints and Drawings, Italy c. 1450–1800. Retrieved on April 27, 2011, from **http://www.iuoart.org/**.

Edge, D., (2001), **The digital imaging of watermarks**, Computing in Musicology, 12, pages 261–274.

Grummer, A., (2009), How to Make Paper: Make a Watermark. Retrieved on April 27, 2011, from **http://www.youtube.com/watch?v=M6h24UXS348**.

Gonzalez , R., Woods , R., and Eddins, S., (2009), **Digital Image Processing Using MATLAB**, (2nd ed.), Upper Saddle River, New Jersey, USA: Pearson Prentice Hall.

Guizar-Sicairos, M., Thurman, S., and Fienup, J., (2008), **Efficient subpixel image registration algorithms**, Optics Letters, vol. 33, No. 2.

Gonzalez , R., and Woods, R., (2008), **Digital Image Processing**, (3rd ed.), Upper Saddle River, New Jersey, USA: Pearson Prentice Hall.

Goshtasby , A., (2005), **2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications**, (1st ed.), Upper Saddle River, New Jersey, USA: Wiley Press.

Gonzalez , R., Woods, R., and Eddins , S., (2004), **Digital Image Processing Using MATLAB**, (1st ed.), Upper Saddle River, New Jersey, USA: Pearson Prentice Hall.

Heawood, E., (1950), **Watermarks mainly of the 17th and 18th centuries**, Paper Publications Society, Hilversum.

Hiary, H., (2008), **Paper-based Watermark Extraction with Image Processing**. Published Doctoral Dissertation, the University of Leeds, Leeds, UK.

Hiary, H., Boyle R., and Ng, K., Digitized Arabic texts from the University of Leeds. Retrieved on April 27, 2011, from **http://www.comp.leeds.ac.uk/arabictexts/**.

Hiary, H., and Ng, K., Paper-based watermark extraction with image processing. Retrieved on April 27, 2011, from **http://icsrim.org.uk/watermark/**.

Haupt, W., (1881), **Wasserzeichenwiedergabe in schwierigen Fallen**, Restauro, 87:38-43.

Koninklijke Bibliotheek - National library of the Netherlands: Watermarks in incunabula printed in the low countries (WILC). Retrieved on April 27, 2011, from **http://watermark.kb.nl/**.

Lewis, J., (1995), **Fast normalized cross-correlation**, Proc. Vision Interface, Quebec City, Canada, 120-123.

LIMA: Literary Manuscript Analysis. Retrieved on April 27, 2011, from **http://www2.warwick.ac.uk/fac/arts/ren/**.

The MathWorks Incorporated, MATLAB software - The Language of Technical Computing. Retrieved on April 27, 2011, from **http://www.mathworks.com/products/matlab/**.

Moschini, D., (2000), **La marca d’acqua: A system for the digital recording of watermarks**. In Puzzles in Paper: Concepts in Historical Watermarks (Papers from the 1996 International Conference on Watermarks at Roanoke, Virginia), 187–192, Oak Knoll Books and the British Library, Newcastle.

Neuheuser, H., Märgner, V., and Meinlschmidt, P., (2005),**Wasserzeichendarstellung mit Hilfe der Thermographie**. ABI TECHNIK 25(4), 266–278.

Otsu, N., (1979), **A threshold selection method from gray-level histograms**, IEEE Transactions on Systems, Man and Cybernetics, 9(1), 62–66.

Special Collections at Leeds University Library. Retrieved on April 27, 2011, from **http://www.leeds.ac.uk/library/spcoll/index.htm**.

Schoonover, D., (1987), **Techniques of reproducingwatermarks: a practical introduction. In: Essays in Paper Analysis**, Folger Shakespeare Library, Washington, pp. 154–167.

Sonka, M., Hlavac , V., and Boyle, R., (2008), **Image Processing, Analysis, and Machine Vision**, (3rd ed.), USA: Thomson.

Stewart, D., Scharf, R., and Arney, J., (1995), **Techniques for digital image capture of watermarks**, Imaging Sci. Technol. 39(3), 261– 267.

Tonazzini, A., and Bianco, G., and Salerno, E., (2009), **Registration and enhancement of double sided degraded manuscripts acquired in multispectral modality**, IEEE 10th International Conference on Document Analysis and Recognition.

Wikipedia, Top-down and bottom-up design. Retrieved on April 27, 2011, from **http://en.wikipedia.org/wiki/Top-down_approach**.

Wang, J., and Tan, C., (2010), **Non-Rigid Image Registration for Historical Manuscript Restoration**, IEEE International Conference on Pattern Recognition.

Wang, J., Brown, M., and Tan, C., (2009), **Automatic Corresponding Control Points Selection for Historical Document Image Registration**, IEEE 10th International Conference on Document Analysis and Recognition.

Wang, J., Brown, M., and Tan, C., (2008), **Accurate Alignment of Double-sided Manuscripts for Bleed-through Removal**, IEEE 8th IAPR International Workshop on Document Analysis Systems (DAS'8), 69–75.

Whelan, P., Soille, P., Drimbarean, A., (2001), **Real-time registration of paper watermarks**, Real Time Imaging 7(4), 367–380.

Wolberg, G., and Zokai, S., (2000), **Robust image registration using log-polar transform**, Proc. of IEEE Intl. Conf. on Image Processing.

Zitova, B., and Flusser, J., (2003), **Image registration methods: A survey**, Image and Vision Computing, vol. 21, 997-1000.

Zamperoni, P., (1989), **Wasserzeichenextraktion aus digitalisierten Bildern mit Methoden der digitalen Bildsignalverarbeitung**. Das Papier, 43(4), 133-143.

Appendix A

Sample Test Data

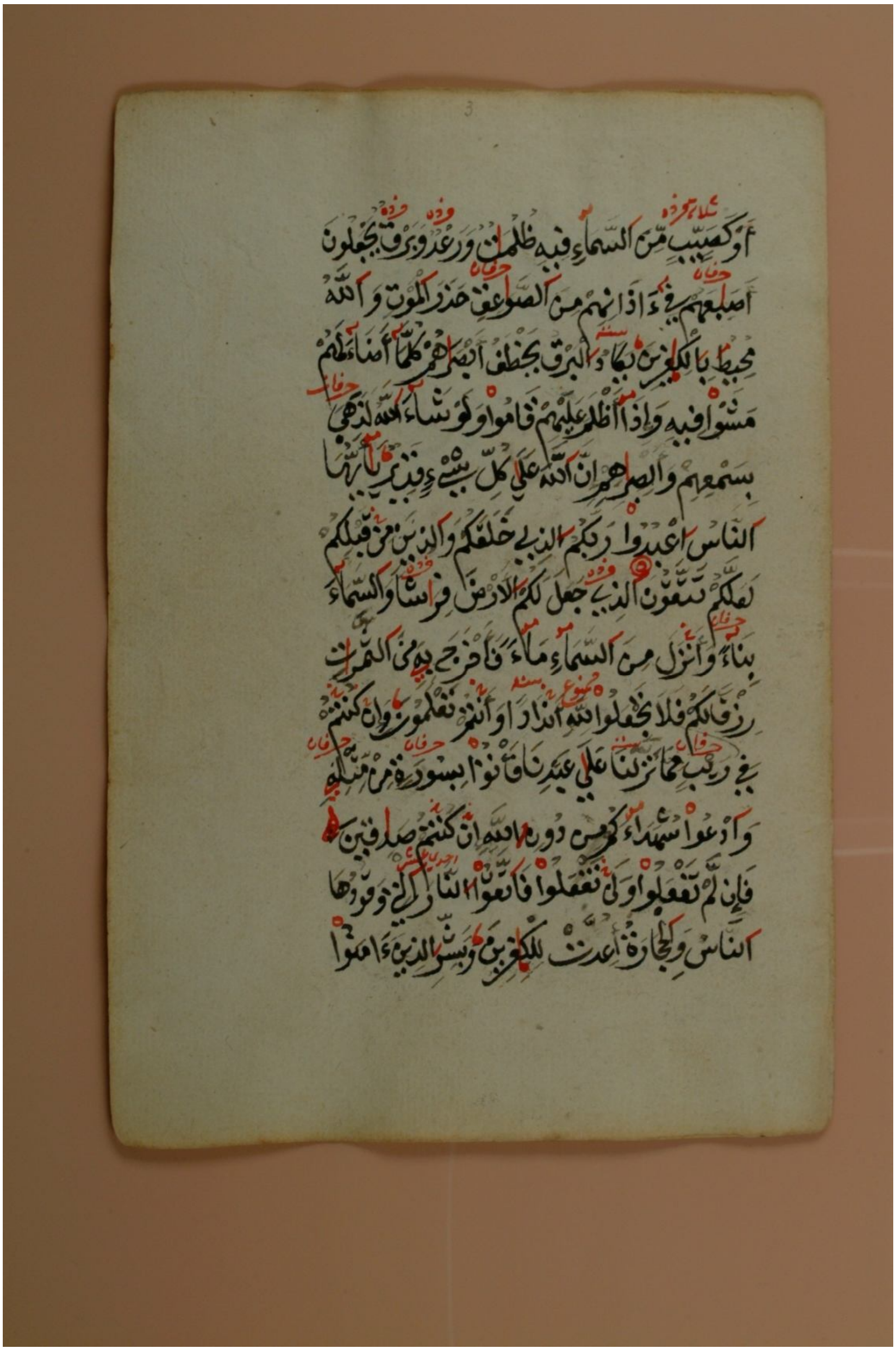


Figure A.1: Reflected recto sample image of page 3 of the 'Mahdiyya' copy of the Qur'an.

Figure A.2: Transmitted recto sample image of page 3 of the 'Mahdiyya' copy of the Qur'an which contains a part of double-headed eagle watermark.


وعملوا الصلحت أن لهم جنت تجري من تحتها الأنهر كلما رزقوا
منها من ثمرة رزقا قالوا هذا الذي رزقنا من قبل
وأتوا به متشبها ولهم فيها أزوج مطهرة وهم فيها خلدون
إن الله لا يستحي أن يضرب مثلا ما بعوضة
فما فوقها فأما الذين ءامنوا فيعلمون أنه الحق من ربهم
وأما الذين كفروا فيقولون ماذا أراد الله بهذا مثلا
يضل به كثيرا ويهدي به كثيرا وما يضل به إلا الفسقين
الذين ينقضون عهد الله من بعد ميثقه ويقطعون
ما أمر الله به أن يوصل ويفسدون في الأرض أولئك هم
الخسرون كيف تكفرون بالله وكنتم أموتا فأحيكم
ثم يميتكم ثم يحييكم ثم إليه ترجعون هو الذي خلق لكم
ما في الأرض جميعا ثم استوى إلى السماء فسوىهن سبع
سموت وهو بكل شيء عليم وإذ قال ربك للملئكة إني
جاعل

Figure A.3: Reflected verso sample image of page 3 the 'Mahdiyya' copy of the Qur'an.

Figure A.4: Transmitted recto sample image of page 5 of the 'Mahdiyya' copy of the Qur'an which contains a part of moonface-within-shield countermark.

Appendix B

Sample Output

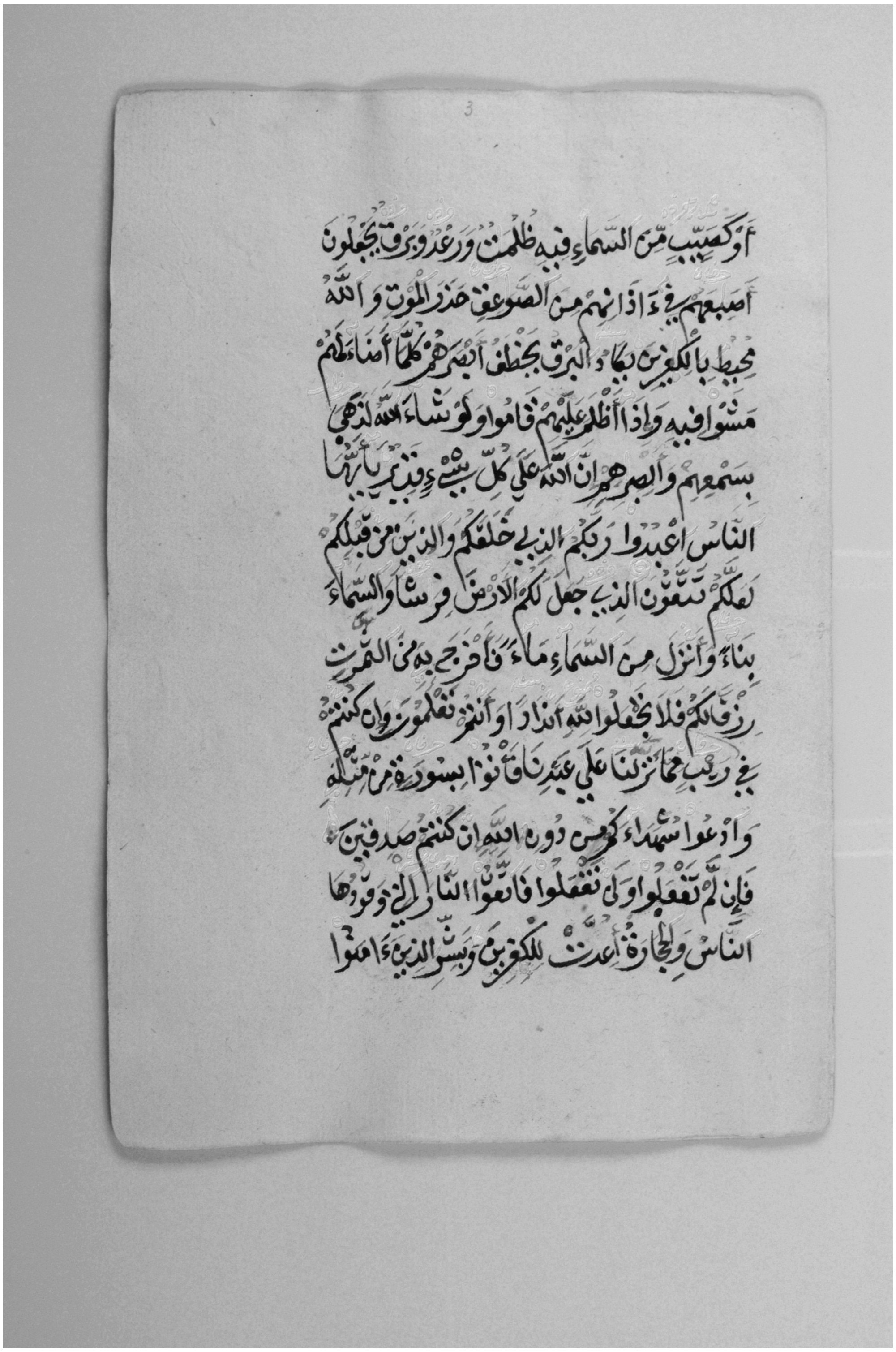
أو كصيب من السماء فيه ظلمت ورعد وبرق يجعلون
أصبعهم في ءاذانهم من الصوعق حذر الموت والله
محيط بالكفرين يكاد البرق يخطف أبصرهم كلما أضاء لهم
مشوا فيه وإذا أظلم عليهم قاموا ولو شاء الله لذهب
بسمعهم وأبصرهم إن الله على كل شيء قدير يأيها
الناس اعبدوا ربكم الذي خلقكم والذين من قبلكم
لعلكم تتقون الذي جعل لكم الأرض فرشا والسماء
بناء وأنزل من السماء ماء فأخرج به من الثمرت
رزقا لكم فلا تجعلوا لله أندادا وأنتم تعلمون وإن كنتم
في ريب مما نزلنا على عبدنا فأتوا بسورة من مثله
وادعوا شهداءكم من دون الله إن كنتم صدقين
فإن لم تفعلوا ولن تفعلوا فاتقوا النار التي وقودها
الناس والحجارة أعدت للكفرين وبشر الذين ءامنوا

Figure B.1: Highlighted RR image.

Figure B.2: Highlighted TR image.


وعملوا الصلحت أن لهم جنت تجري من تحتها الأنهر كلما رزقوا
منها من ثمرة رزقا قالوا هذا الذي رزقنا من قبل
وأتوا به متشبها ولهم فيها أزوج مطهرة وهم فيها خلدون
إن الله لا يستحي أن يضرب مثلا ما بعوضة
فما فوقها فأما الذين ءامنوا فيعلمون أنه الحق من ربهم
وأما الذين كفروا فيقولون ماذا أراد الله بهذا مثلا
يضل به كثيرا ويهدي به كثيرا وما يضل به إلا الفسقين
الذين ينقضون عهد الله من بعد ميثقه ويقطعون
ما أمر الله به أن يوصل ويفسدون في الأرض أولئك هم
الخسرون كيف تكفرون بالله وكنتم أموتا فأحيكم
ثم يميتكم ثم يحييكم ثم إليه ترجعون هو الذي خلق لكم
ما في الأرض جميعا ثم استوى إلى السماء فسوىهن سبع
سموت وهو بكل شيء عليم وإذ قال ربك للملئكة إني
جاعل

Figure B.3: Highlighted RV image.



3

أو كصيب من السماء فيه ظلمت ورعد وبرق يجعلون
أصبعهم في ءاذانهم من الصوعق حذر الموت والله
محيط بالكفرين يكاد البرق يخطف أبصرهم كلما أضاء لهم
مشوا فيه وإذا أظلم عليهم قاموا ولو شاء الله لذهب
بسمعهم وأبصرهم إن الله على كل شيء قدير يأيها
الناس اعبدوا ربكم الذي خلقكم والذين من قبلكم
لعلكم تتقون الذي جعل لكم الأرض فرشا والسماء
بناء وأنزل من السماء ماء فأخرج به من الثمرت
رزقا لكم فلا تجعلوا لله أندادا وأنتم تعلمون وإن كنتم
في ريب مما نزلنا على عبدنا فأتوا بسورة من مثله
وادعوا شهداءكم من دون الله إن كنتم صدقين
فإن لم تفعلوا ولن تفعلوا فاتقوا النار التي وقودها
الناس والحجارة أعدت للكفرين وبشر الذين ءامنوا

Figure B.4: Monochromatic RR image.

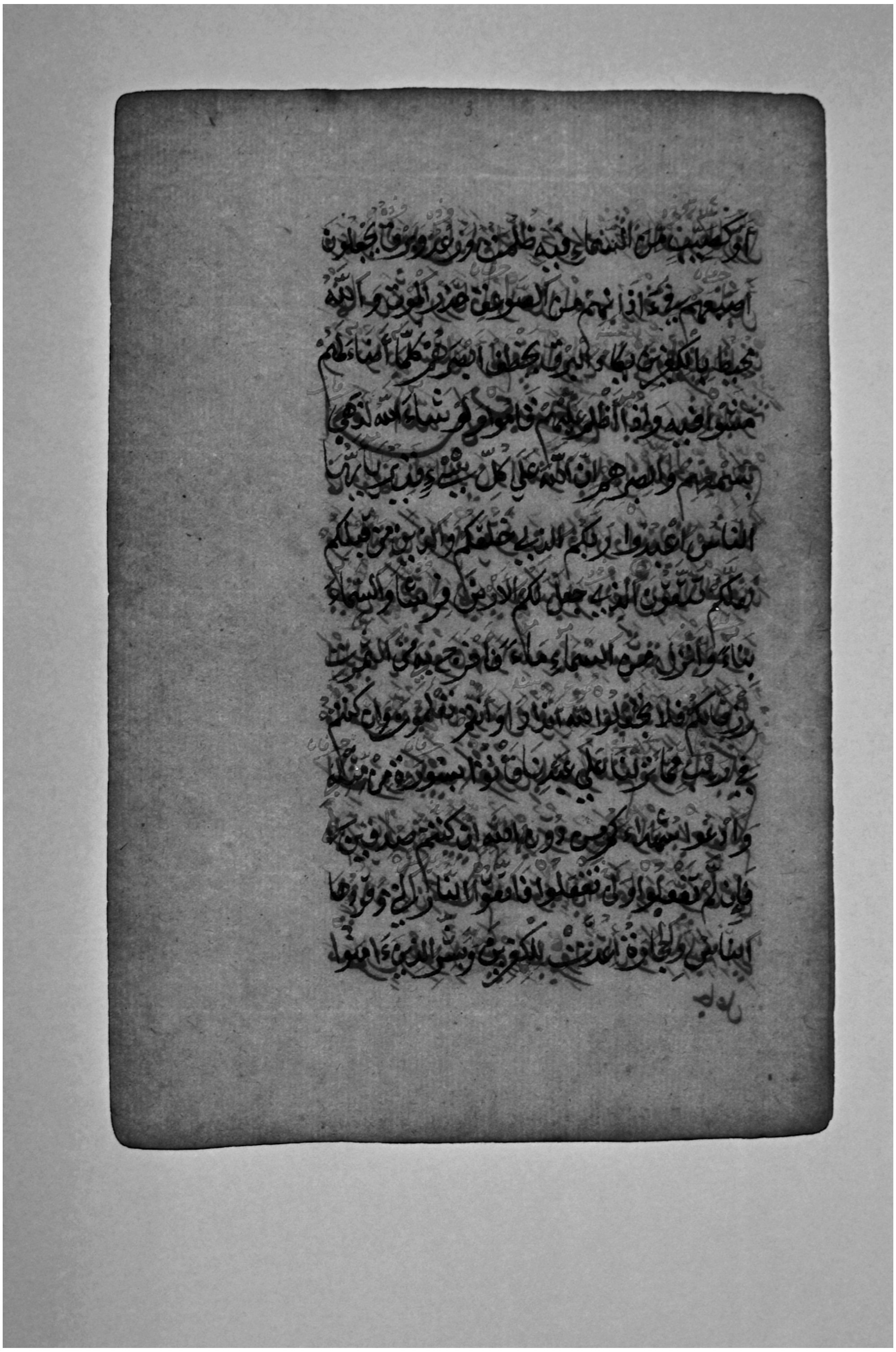
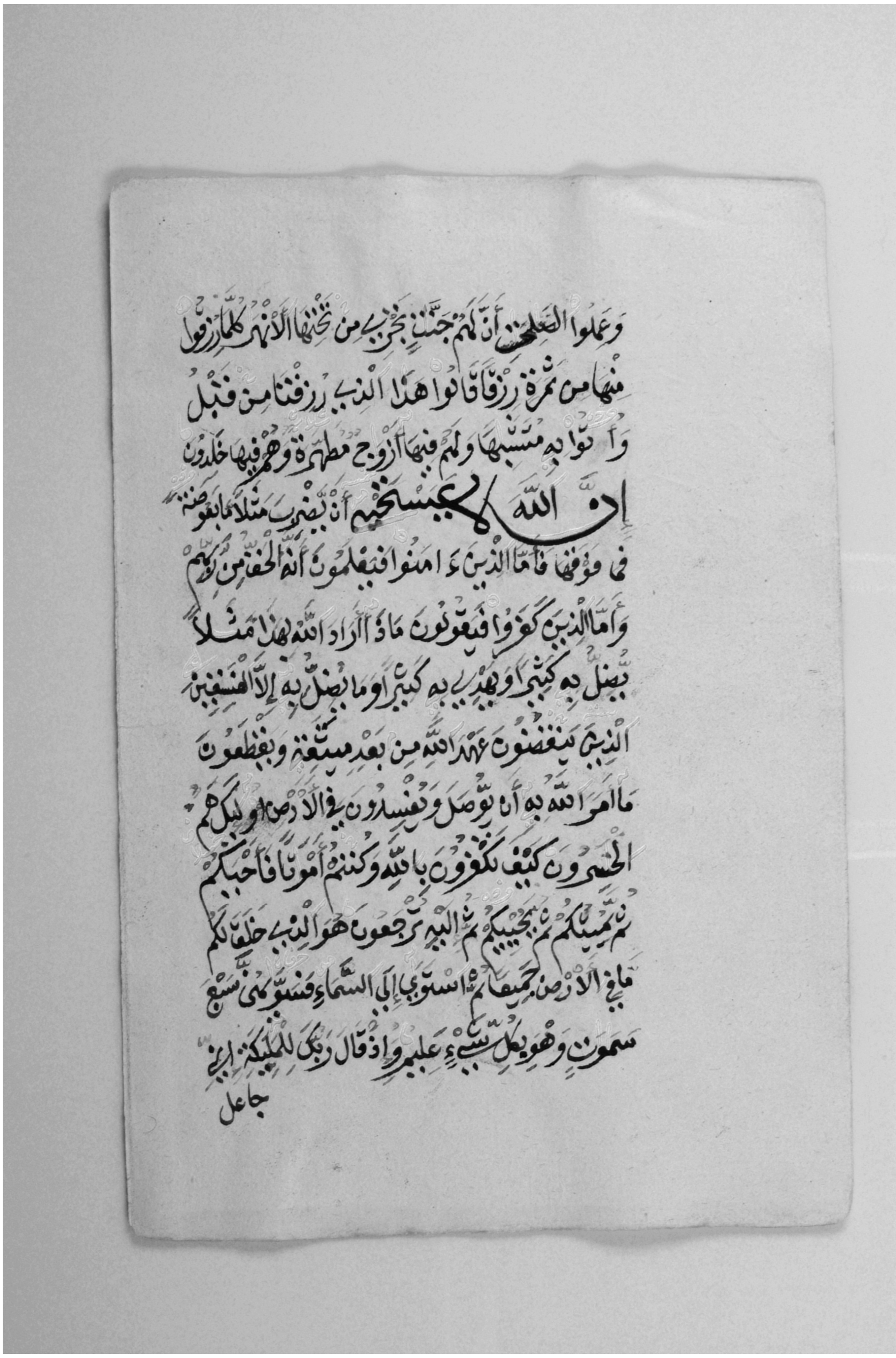


Figure B.5: Monochromatic TR image.



وعملوا الصالحات أن لهم جنات تجري من تحتها الأنهار كلما رزقوا
منها من ثمرة رزقا قالوا هذا الذي رزقنا من قبل
وأتوا به متشابها ولهم فيها أزواج مطهرة وهم فيها خالدون
إن الله لا يستحيي أن يضرب مثلا ما بعوضة
فما فوقها فأما الذين آمنوا فيعلمون أنه الحق من ربهم
وأما الذين كفروا فيقولون ماذا أراد الله بهذا مثلا
يضل به كثيرا ويهدي به كثيرا وما يضل به إلا الفاسقين
الذين ينقضون عهد الله من بعد ميثاقه ويقطعون
ما أمر الله به أن يوصل ويفسدون في الأرض أولئك هم
الخاسرون كيف تكفرون بالله وكنتم أمواتا فأحياكم
ثم يميتكم ثم يحييكم ثم إليه ترجعون هو الذي خلق لكم
ما في الأرض جميعا ثم استوى إلى السماء فسواهن سبع
سموات وهو بكل شيء عليم وإذ قال ربك للملائكة إني
جاعل

Figure B.6: Monochromatic RV image.

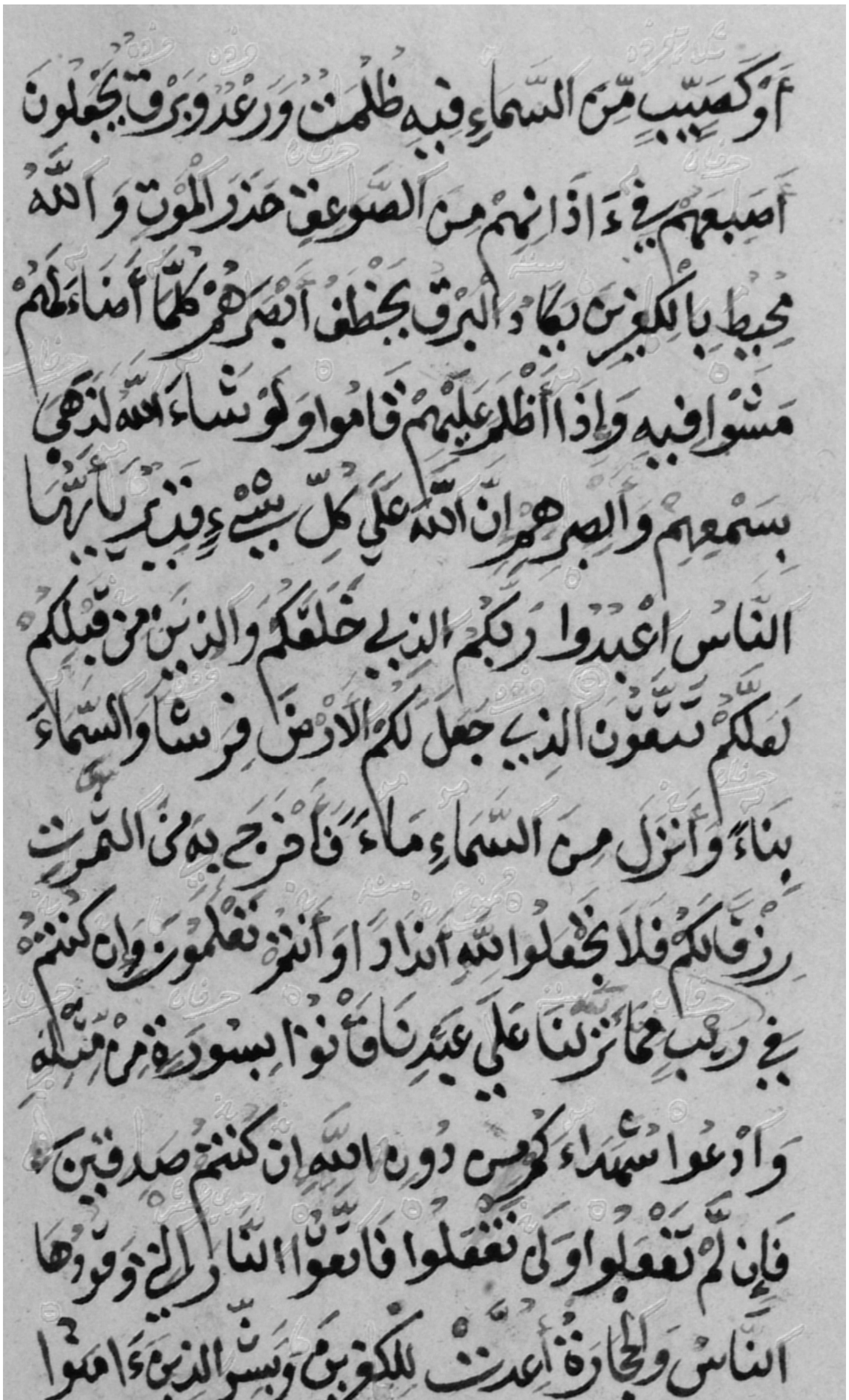

أو كصيب من السماء فيه ظلمت ورعد وبرق يجعلون
أصبعهم في ءاذانهم من الصوعق حذر الموت والله
محيط بالكفرين يكاد البرق يخطف أبصرهم كلما أضاء لهم
مشوا فيه وإذا أظلم عليهم قاموا ولو شاء الله لذهب
بسمعهم وأبصرهم إن الله على كل شيء قدير يأيها
الناس اعبدوا ربكم الذي خلقكم والذين من قبلكم
لعلكم تتقون الذي جعل لكم الأرض فرشا والسماء
بناء وأنزل من السماء ماء فأخرج به من الثمرت
رزقا لكم فلا تجعلوا لله أندادا وأنتم تعلمون وإن كنتم
في ريب مما نزلنا على عبدنا فأتوا بسورة من مثله
وادعوا شهداءكم من دون الله إن كنتم صدقين
فإن لم تفعلوا ولن تفعلوا فاتقوا النار التي وقودها
الناس والحجارة أعدت للكفرين وبشر الذين ءامنوا

Figure B.7: ROI of RR image.

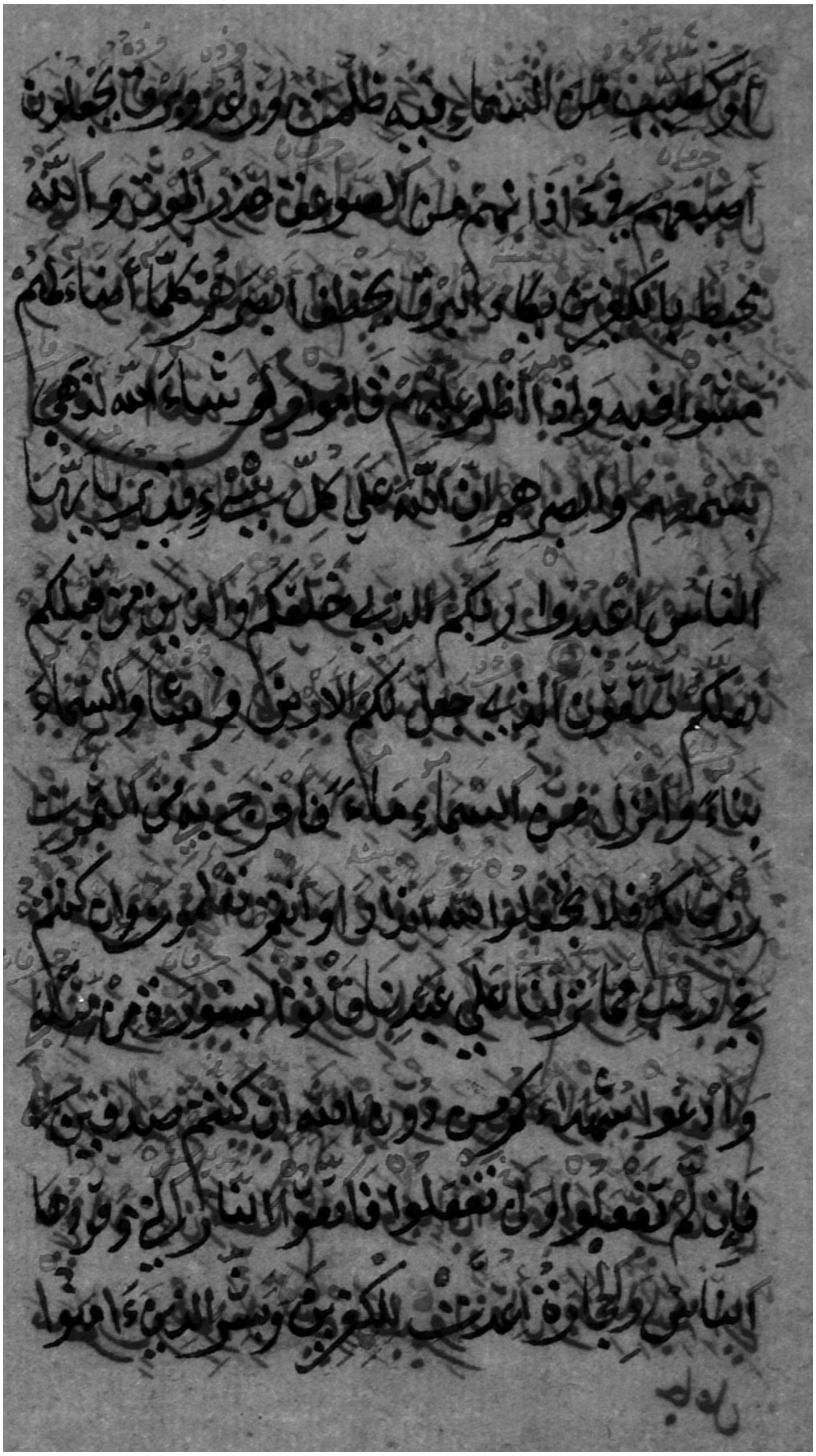

أو كصيب من السماء فيه ظلمت ورعد وبرق يجعلون
أصبعهم في ءاذانهم من الصوعق حذر الموت والله
محيط بالكفرين يكاد البرق يخطف أبصرهم كلما أضاء لهم
مشوا فيه وإذا أظلم عليهم قاموا ولو شاء الله لذهب
بسمعهم وأبصرهم إن الله على كل شيء قدير يأيها
الناس اعبدوا ربكم الذي خلقكم والذين من قبلكم
لعلكم تتقون الذي جعل لكم الأرض فرشا والسماء
بناء وأنزل من السماء ماء فأخرج به من الثمرت
رزقا لكم فلا تجعلوا لله أندادا وأنتم تعلمون وإن كنتم
في ريب مما نزلنا على عبدنا فأتوا بسورة من مثله
وادعوا شهداءكم من دون الله إن كنتم صدقين
فإن لم تفعلوا ولن تفعلوا فاتقوا النار التي وقودها
الناس والحجارة أعدت للكفرين وبشر الذين ءامنوا

Figure B.8: ROI of TR image.

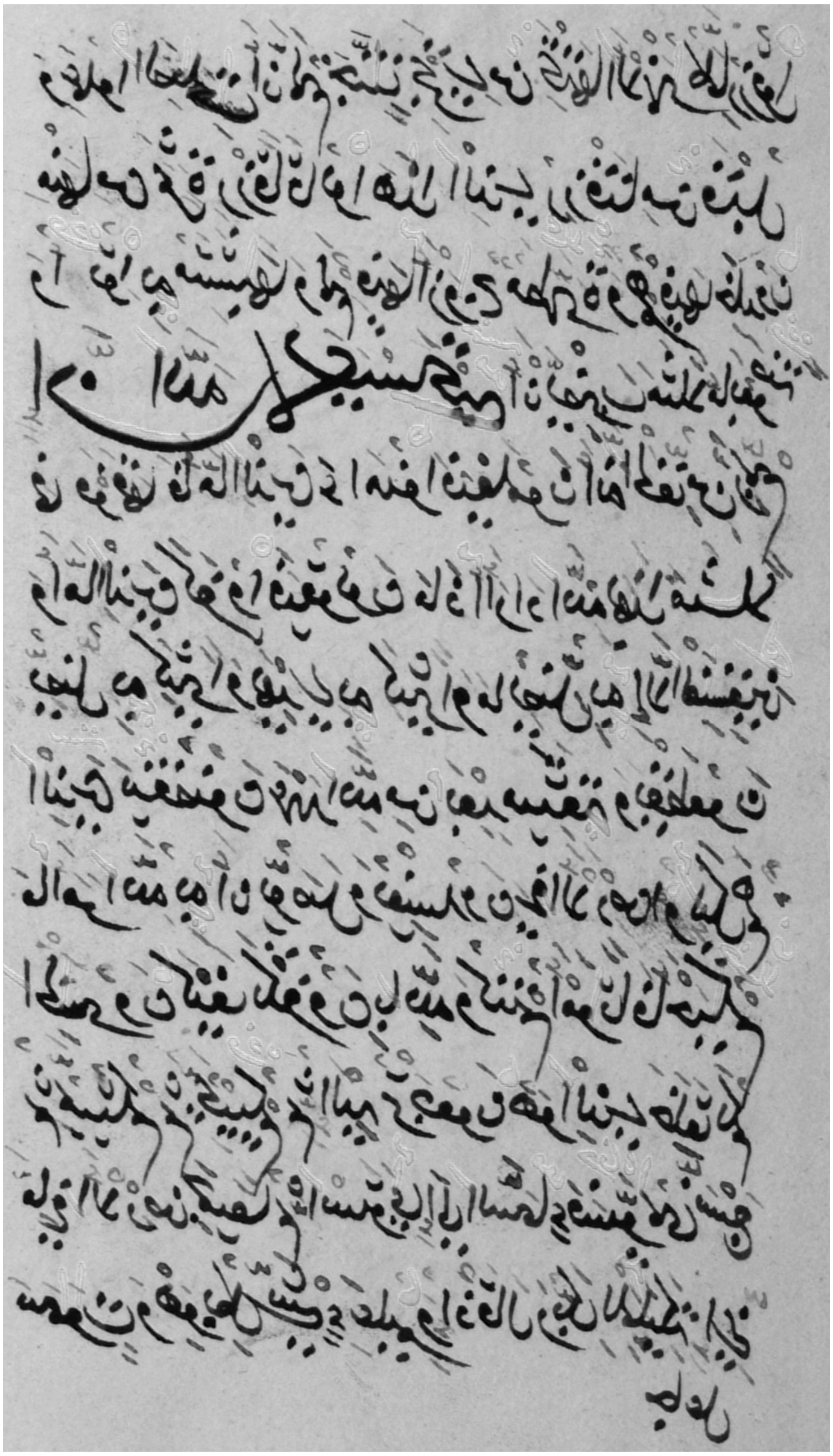


Figure B.9: ROI of RV image (flipped horizontally).

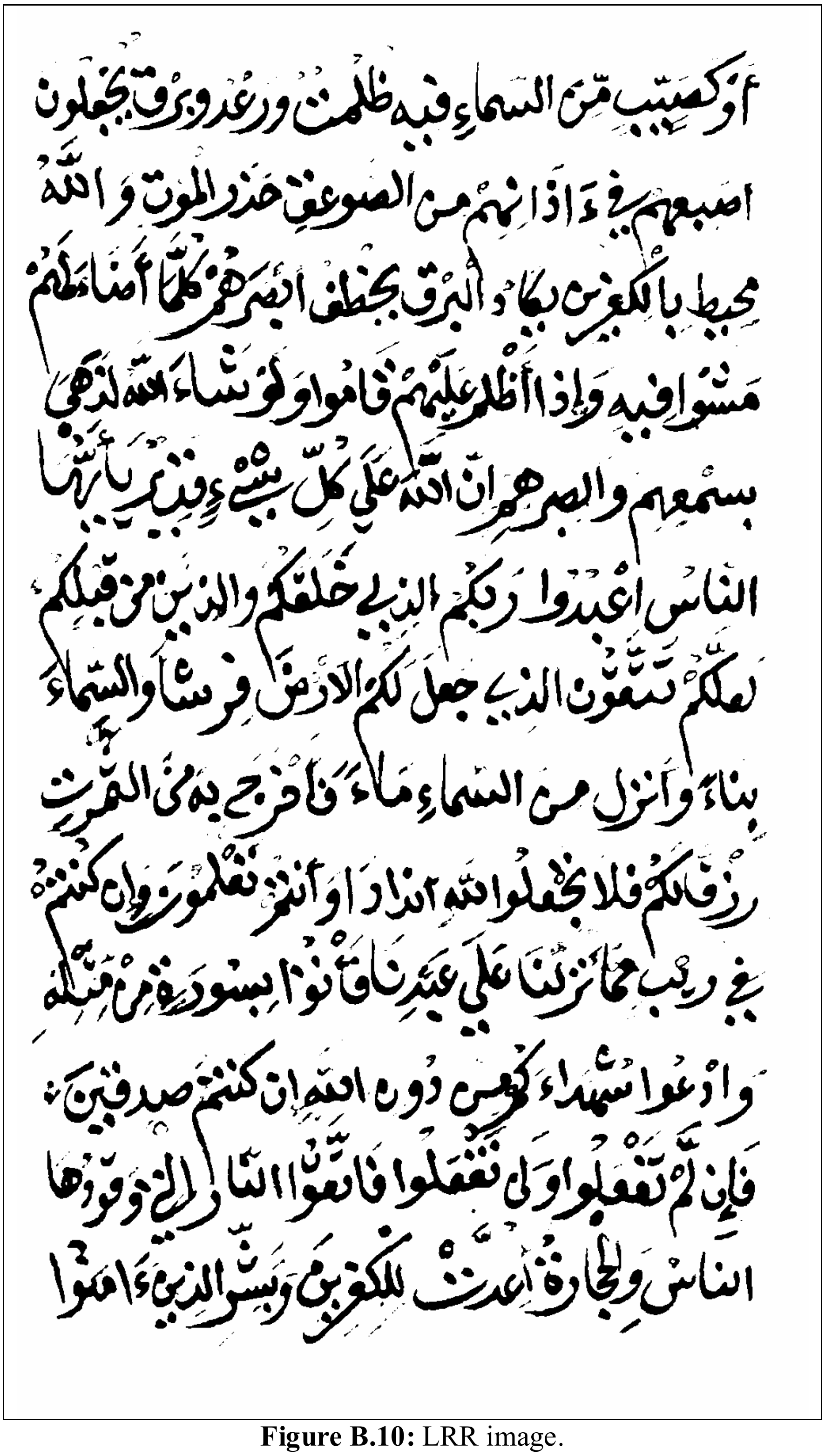

أو كصيب من السماء فيه ظلمت ورعد وبرق يجعلون
اصبعهم في ءاذانهم من الصوعق حذر الموت والله
محيط بالكفرين يكاد البرق يخطف أبصرهم كلما أضاء لهم
مشوا فيه وإذا أظلم عليهم قاموا ولو شاء الله لذهب
بسمعهم وأبصرهم إن الله على كل شيء قدير يأيها
الناس اعبدوا ربكم الذي خلقكم والذين من قبلكم
لعلكم تتقون الذي جعل لكم الأرض فرشا والسماء
بناء وأنزل من السماء ماء فأخرج به من الثمرت
رزقا لكم فلا تجعلوا لله أندادا وأنتم تعلمون وإن كنتم
في ريب مما نزلنا على عبدنا فأتوا بسورة من مثله
وادعوا شهداءكم من دون الله إن كنتم صدقين
فإن لم تفعلوا ولن تفعلوا فاتقوا النار التي وقودها
الناس والحجارة أعدت للكفرين وبشر الذين ءامنوا

Figure B.10: LRR image.

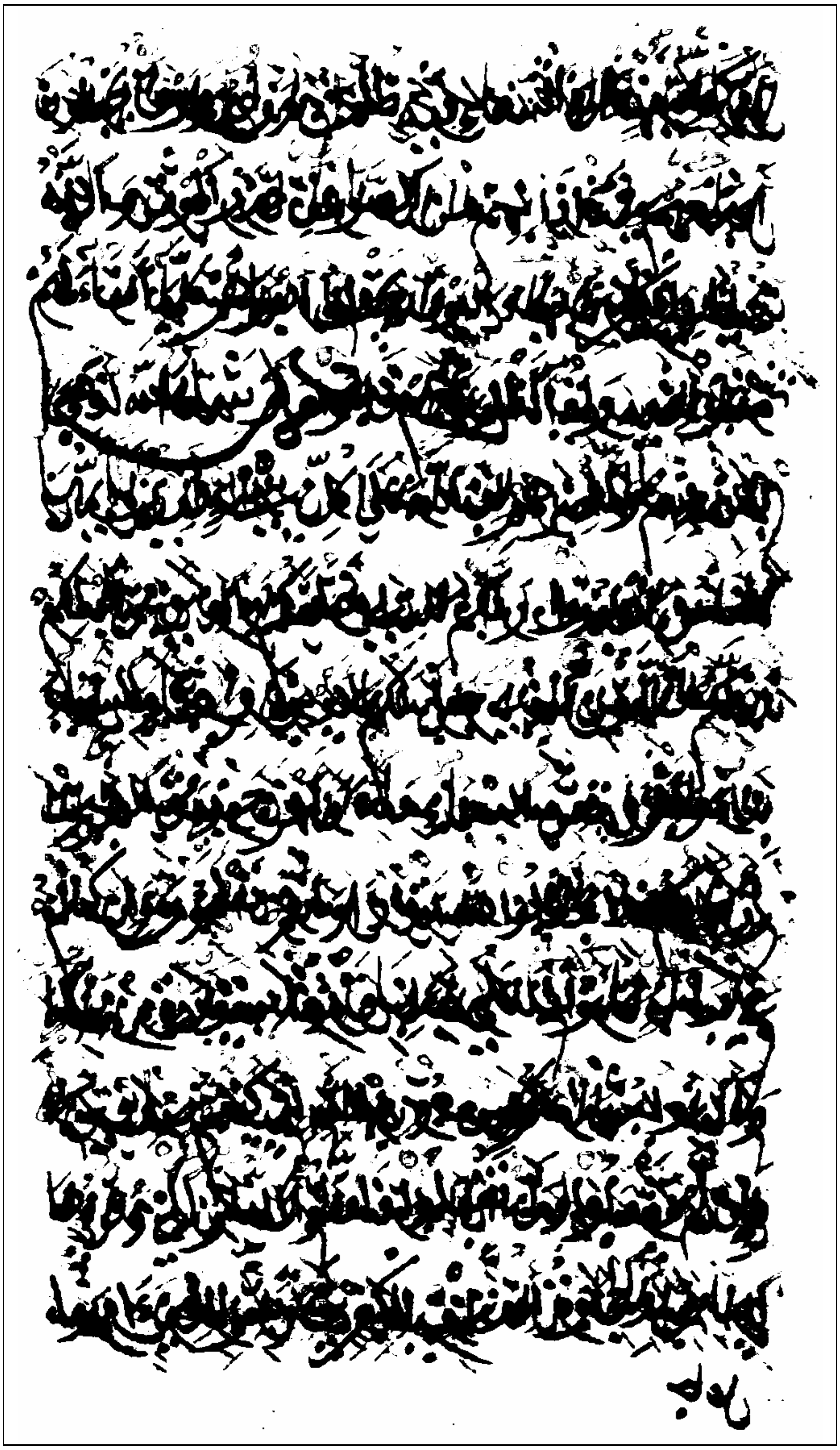


Figure B.11: LTR image.

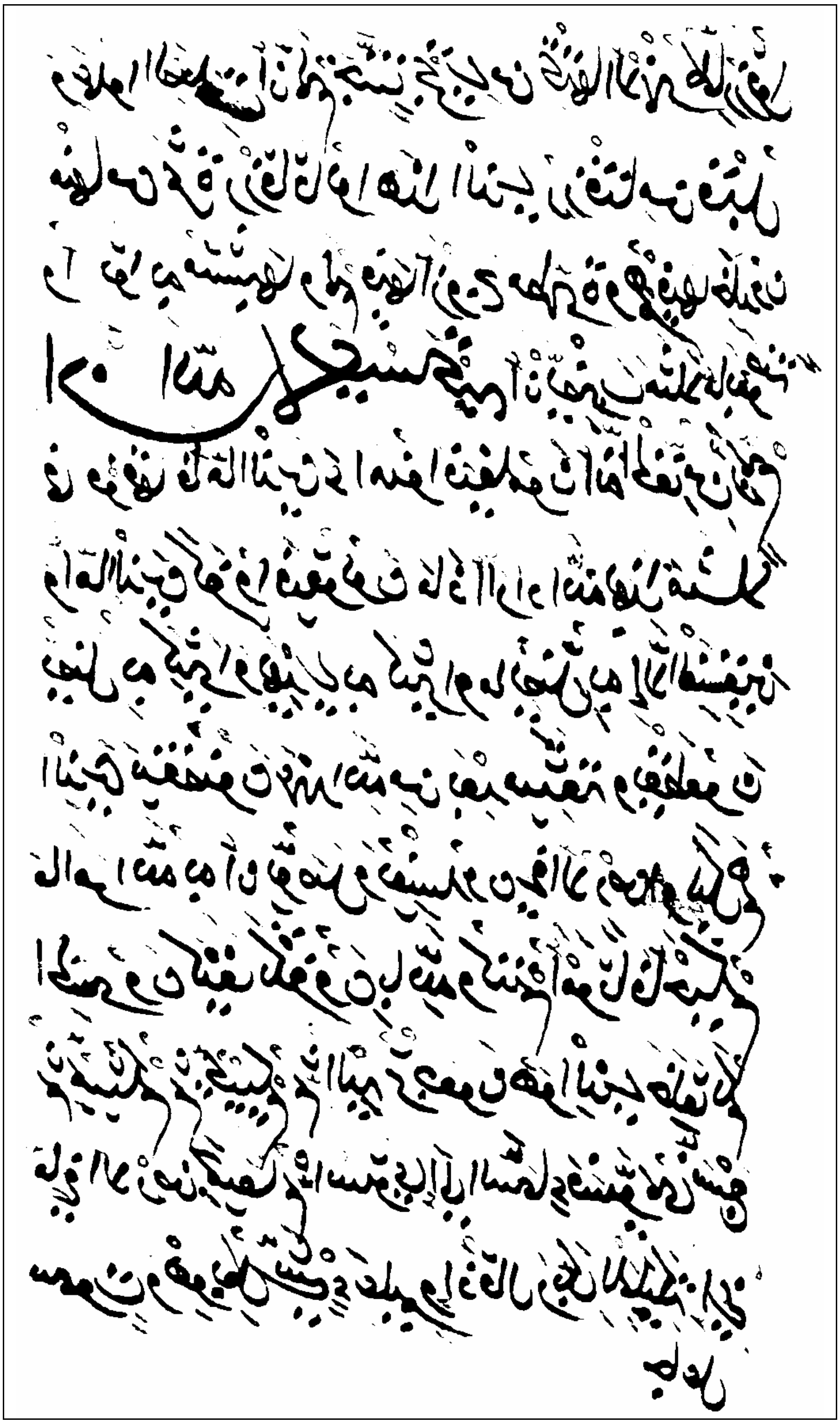


Figure B.12: LRV image (flipped horizontally).

أو كصيب من السماء فيه ظلمت ورعد وبرق يجعلون
اصبعهم في اذانهم من الصوعق حذر الموت والله
محيط بالكفرين يكاد البرق يخطف ابصرهم كلما اضاء لهم
مشوا فيه وإذا أظلم عليهم قاموا ولو شاء الله لذهب
بسمعهم وابصرهم ان الله على كل شيء قدير يايها
الناس اعبدوا ربكم الذي خلقكم والذين من قبلكم
لعلكم تتقون الذي جعل لكم الارض فرشا والسماء
بناء وانزل من السماء ماء فاخرج به من الثمرت
رزقا لكم فلا تجعلوا لله اندادا وانتم تعلمون وان كنتم
في ريب مما نزلنا على عبدنا فاتوا بسورة من مثله
وادعوا شهداءكم من دون الله ان كنتم صدقين
فإن لم تفعلوا ولن تفعلوا فاتقوا النار التي وقودها
الناس والحجارة اعدت للكفرين وبشر الذين ءامنوا

Figure B.13: Padded LRR image.

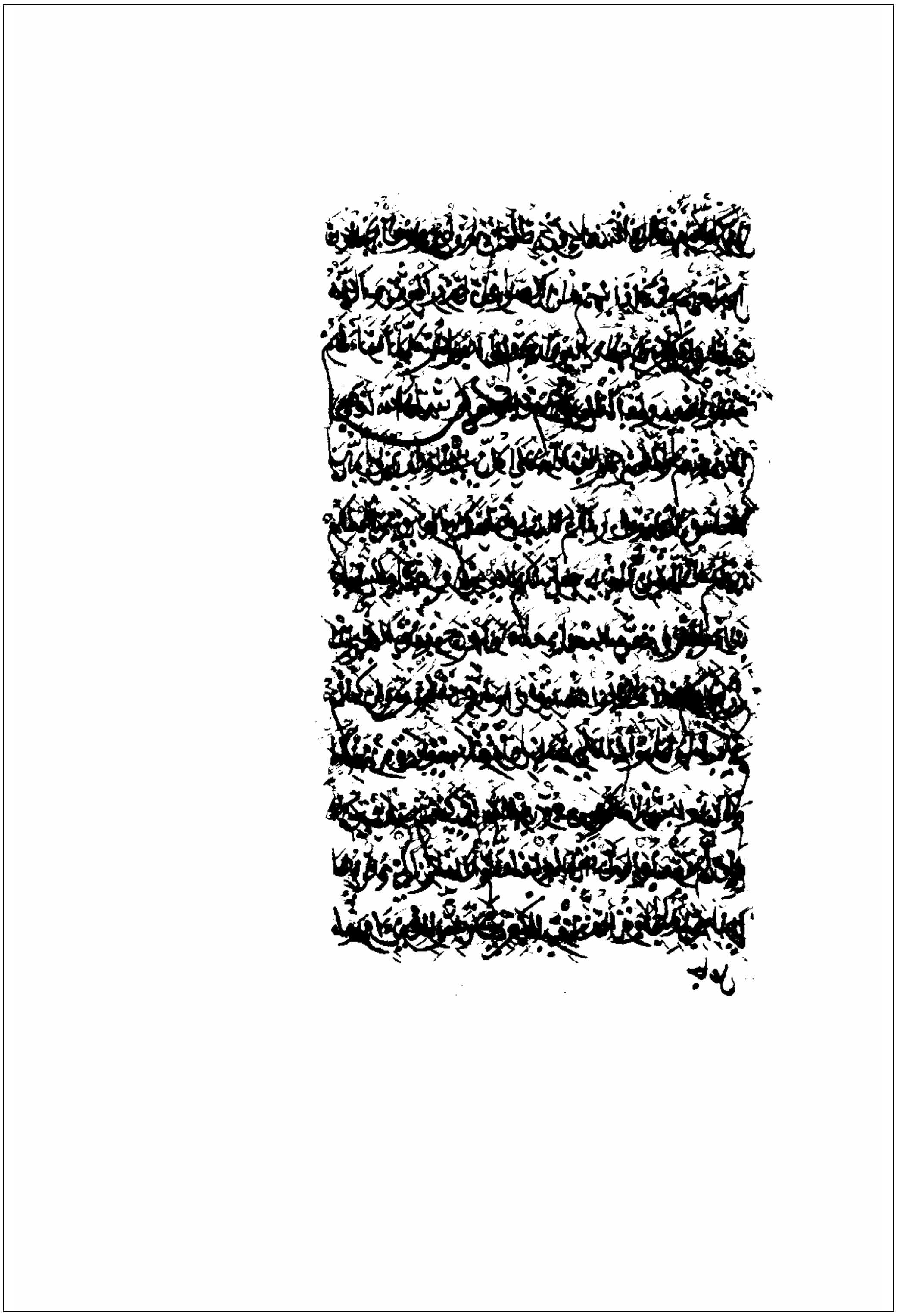


Figure B.14: Padded LTR image.

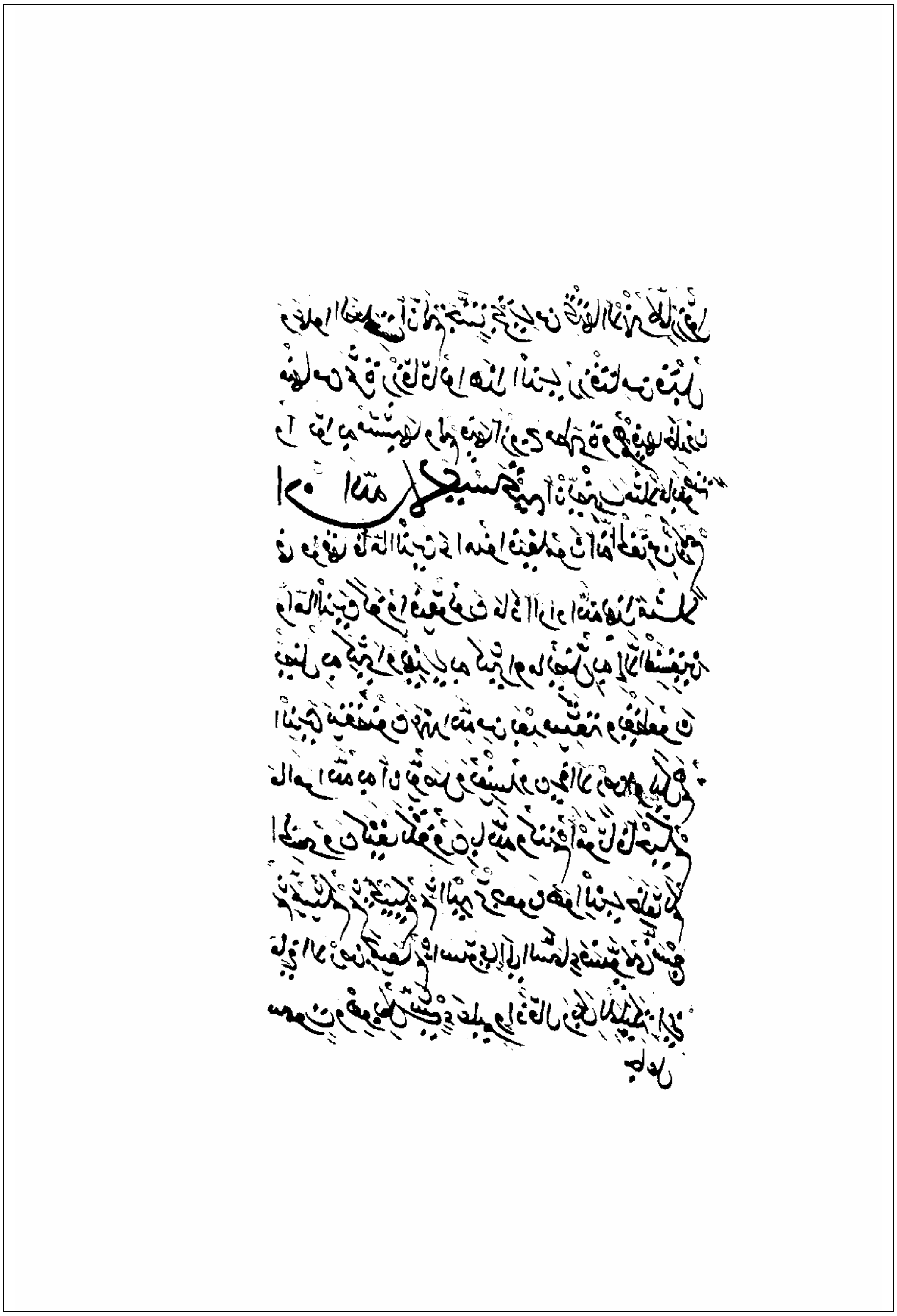


Figure B.15: Padded LRV image.

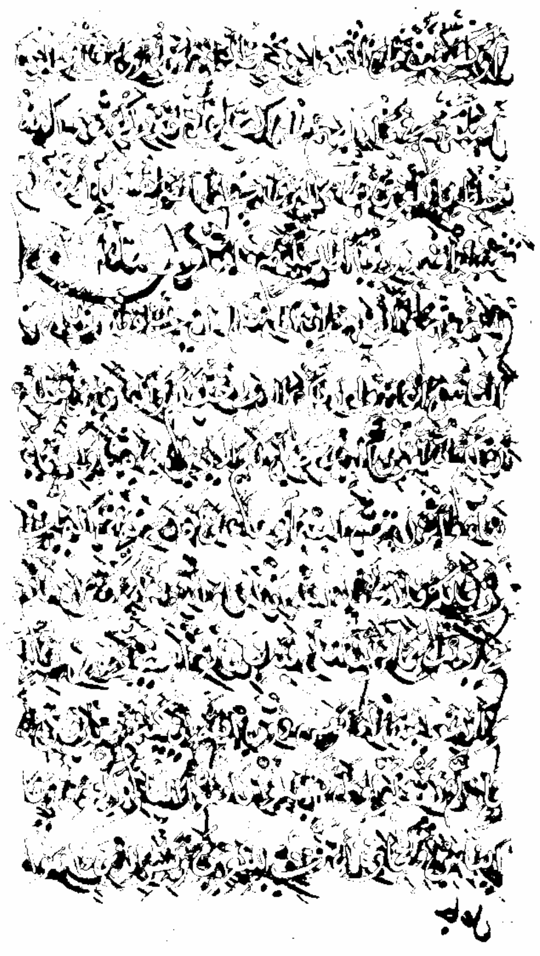


Figure B.16: EVF.

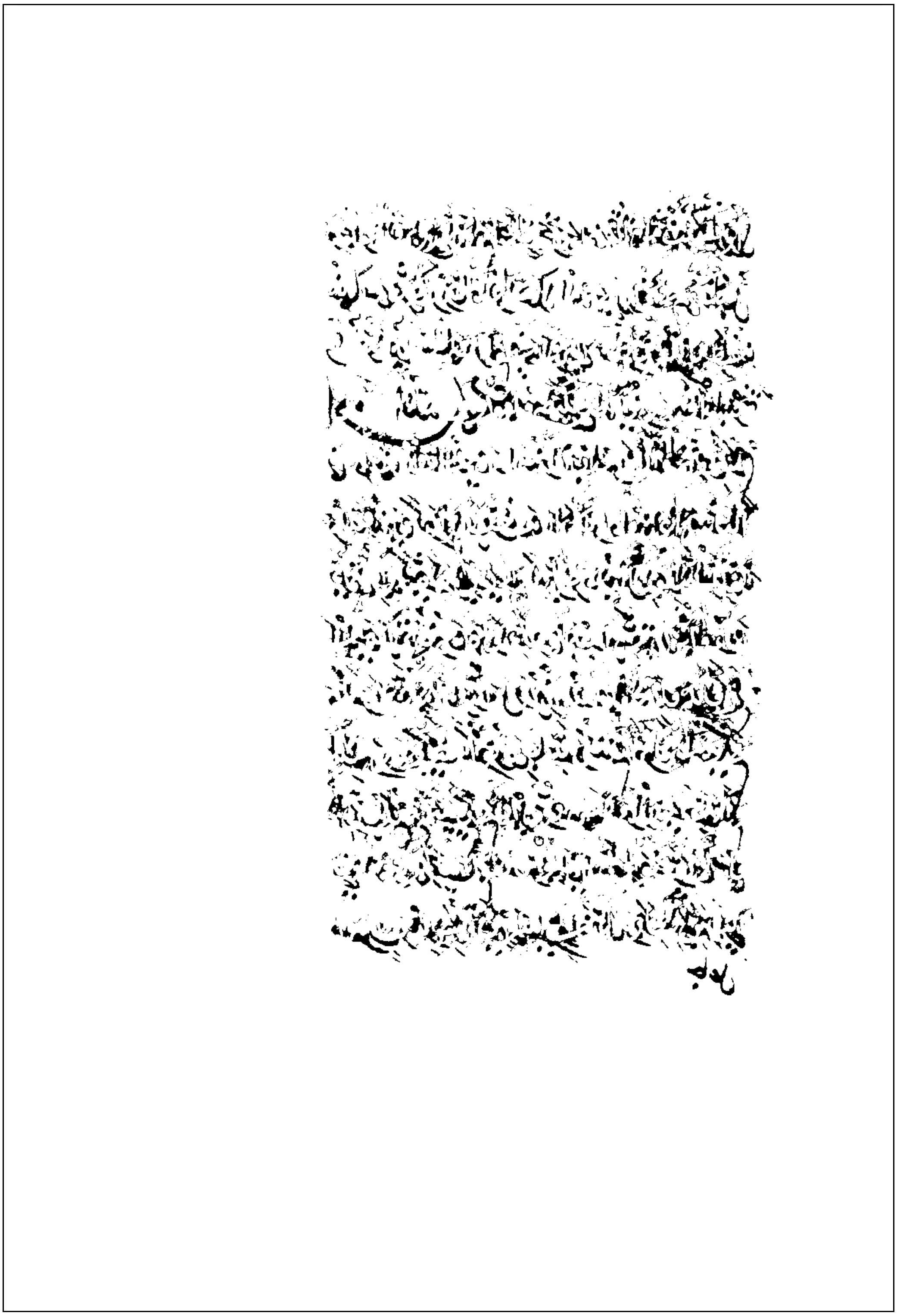


Figure B.17: EVF after erosion.

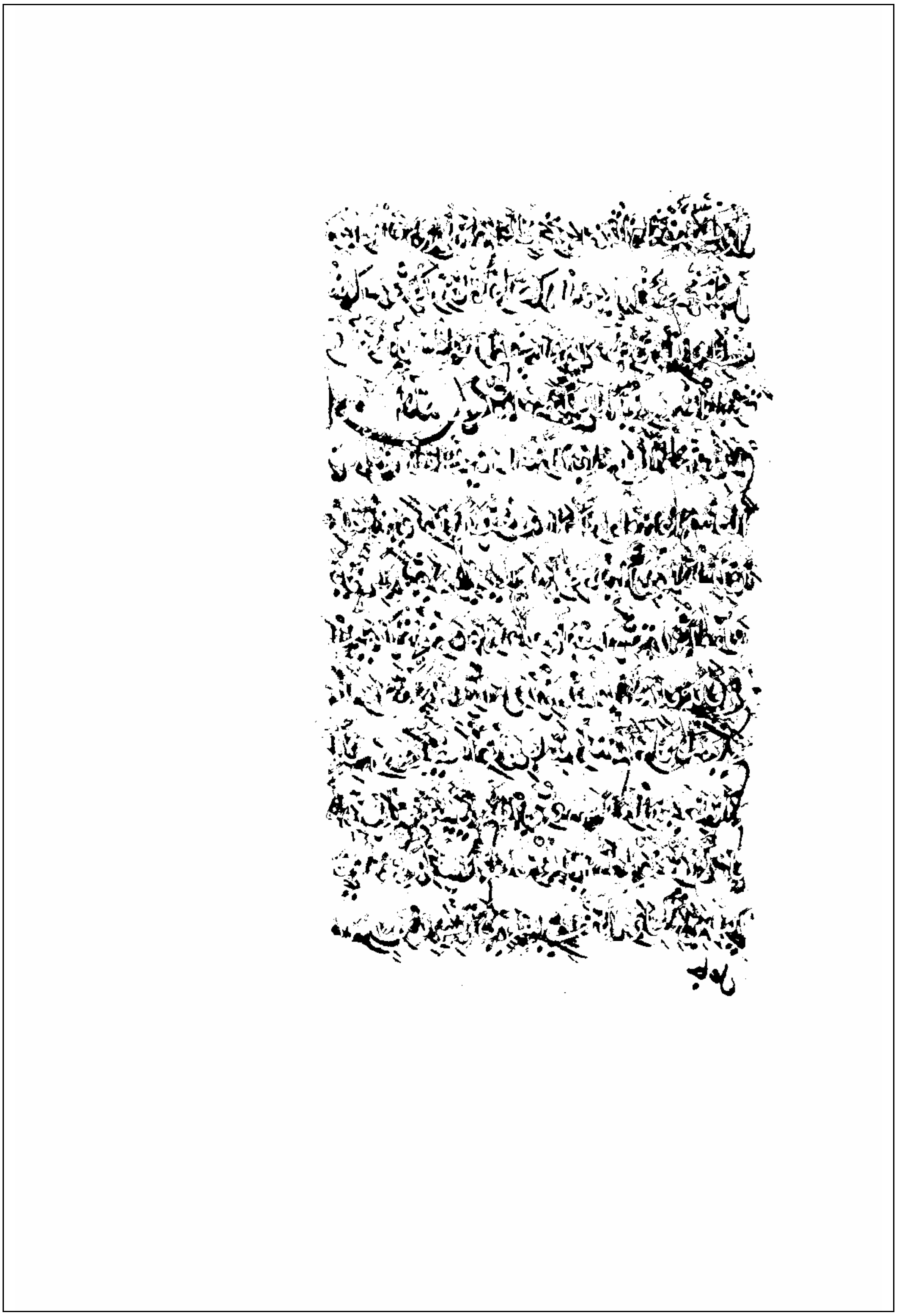


Figure B.18: EVF after dilation.

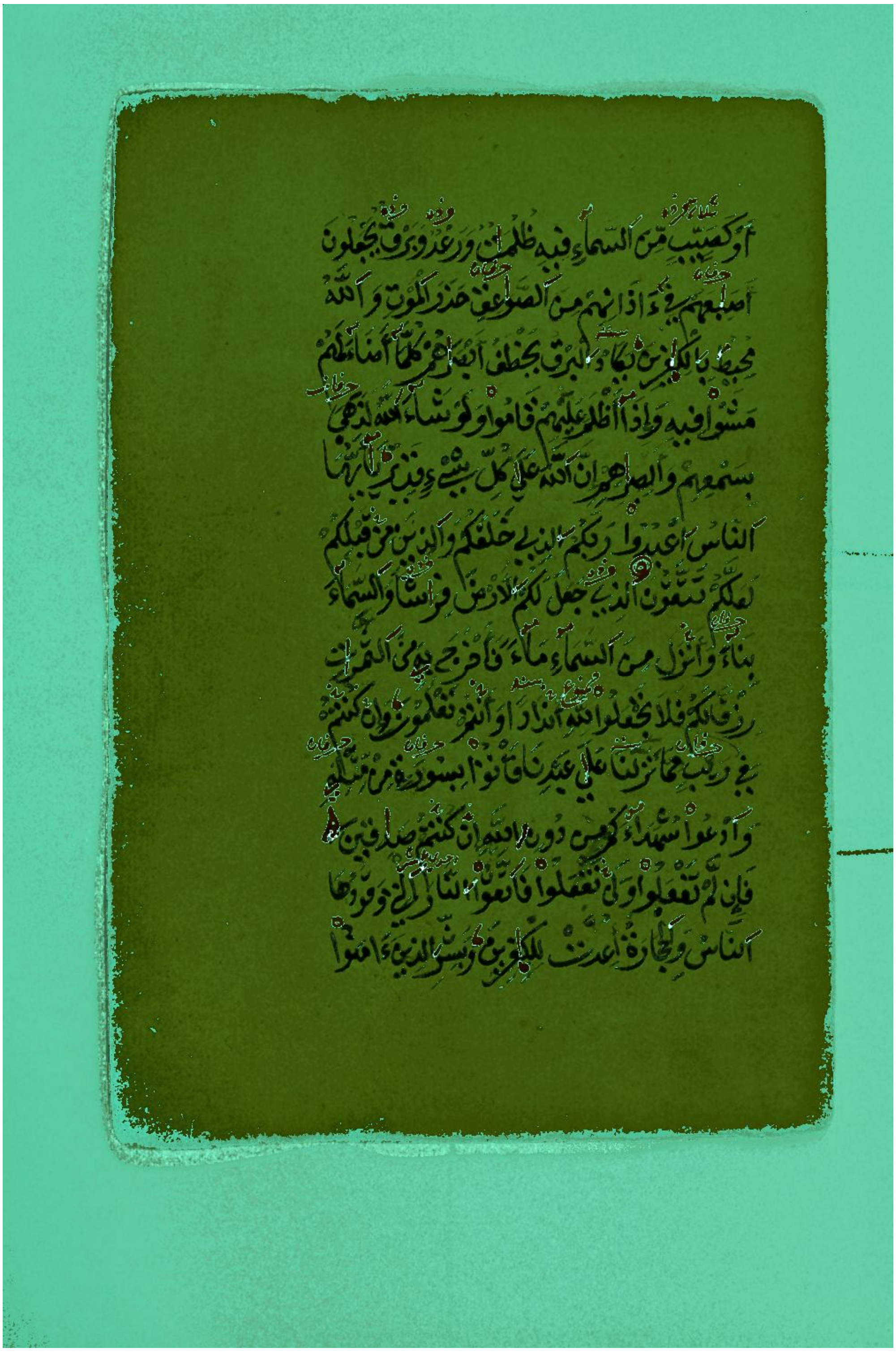


Figure B.19: Reflected recto scan after back-lighting approximation.

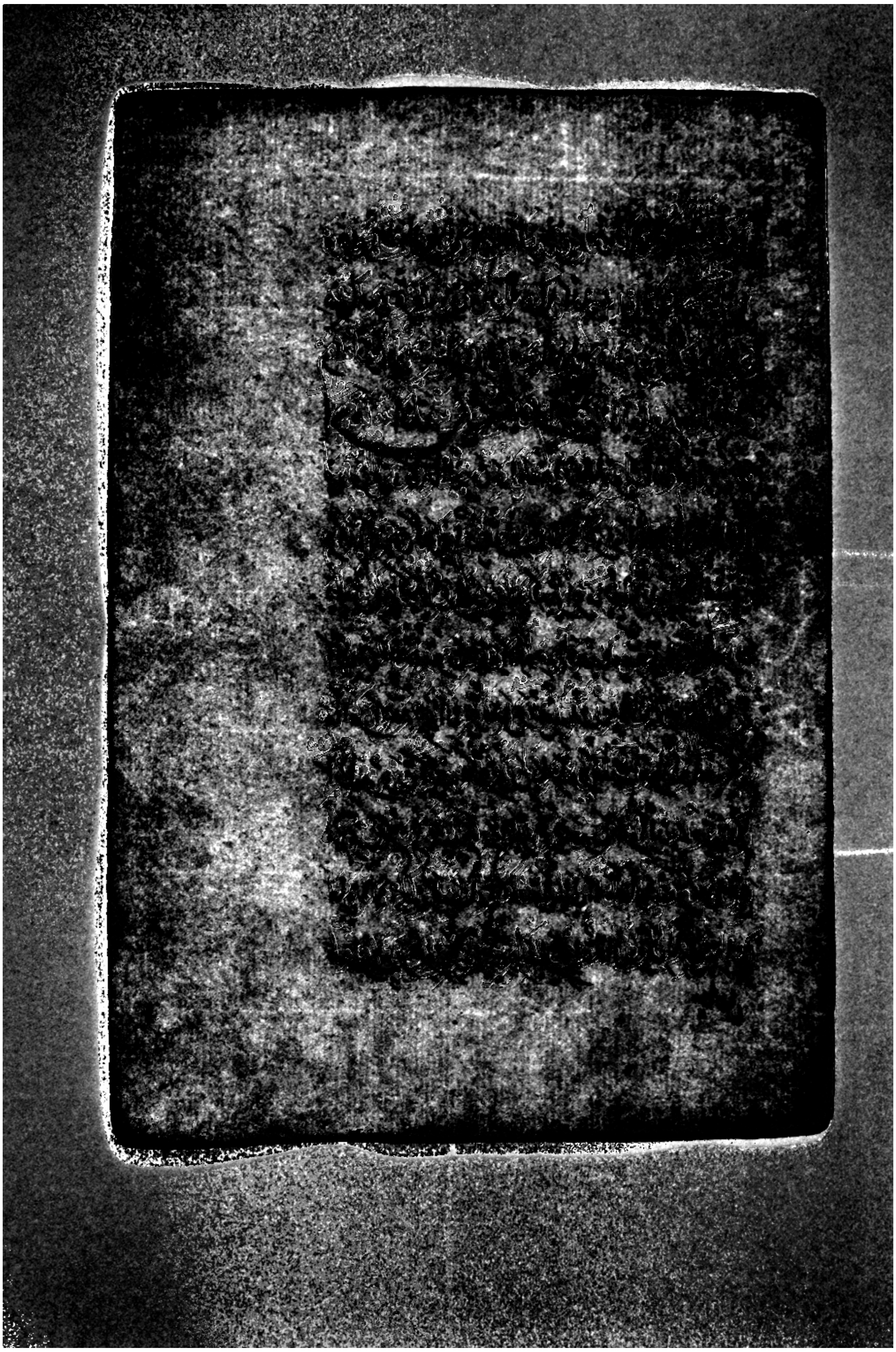


Figure B.20: Result of the differencing operation between approximated reflected recto and its transmitted recto images (enhanced for display).

Figure B.21: Registered reflected verso scan after back-lighting approximation.

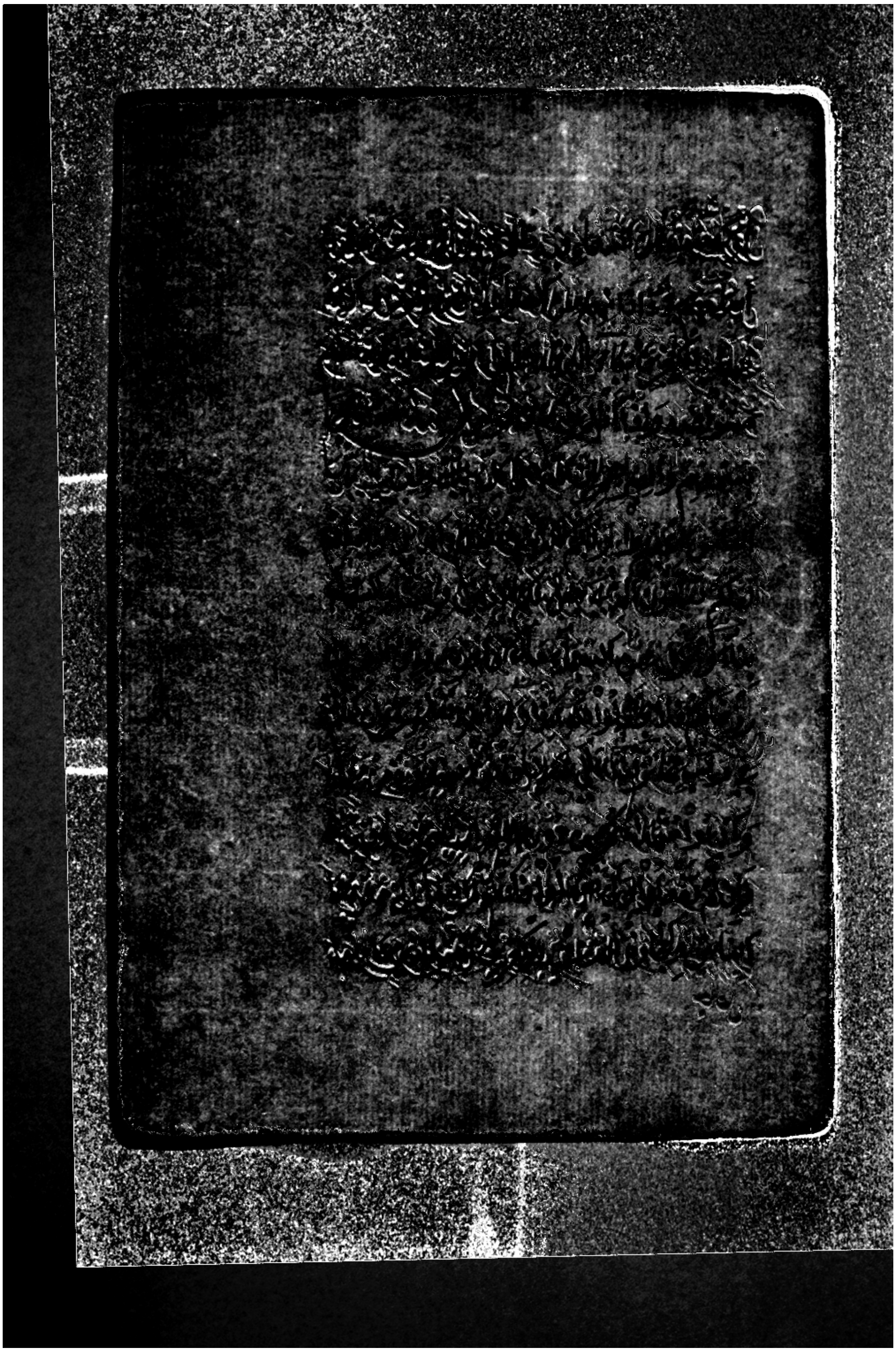


Figure B.22: Result of the differencing operation between transformed registered reflected verso and transmitted recto images (enhanced for display).

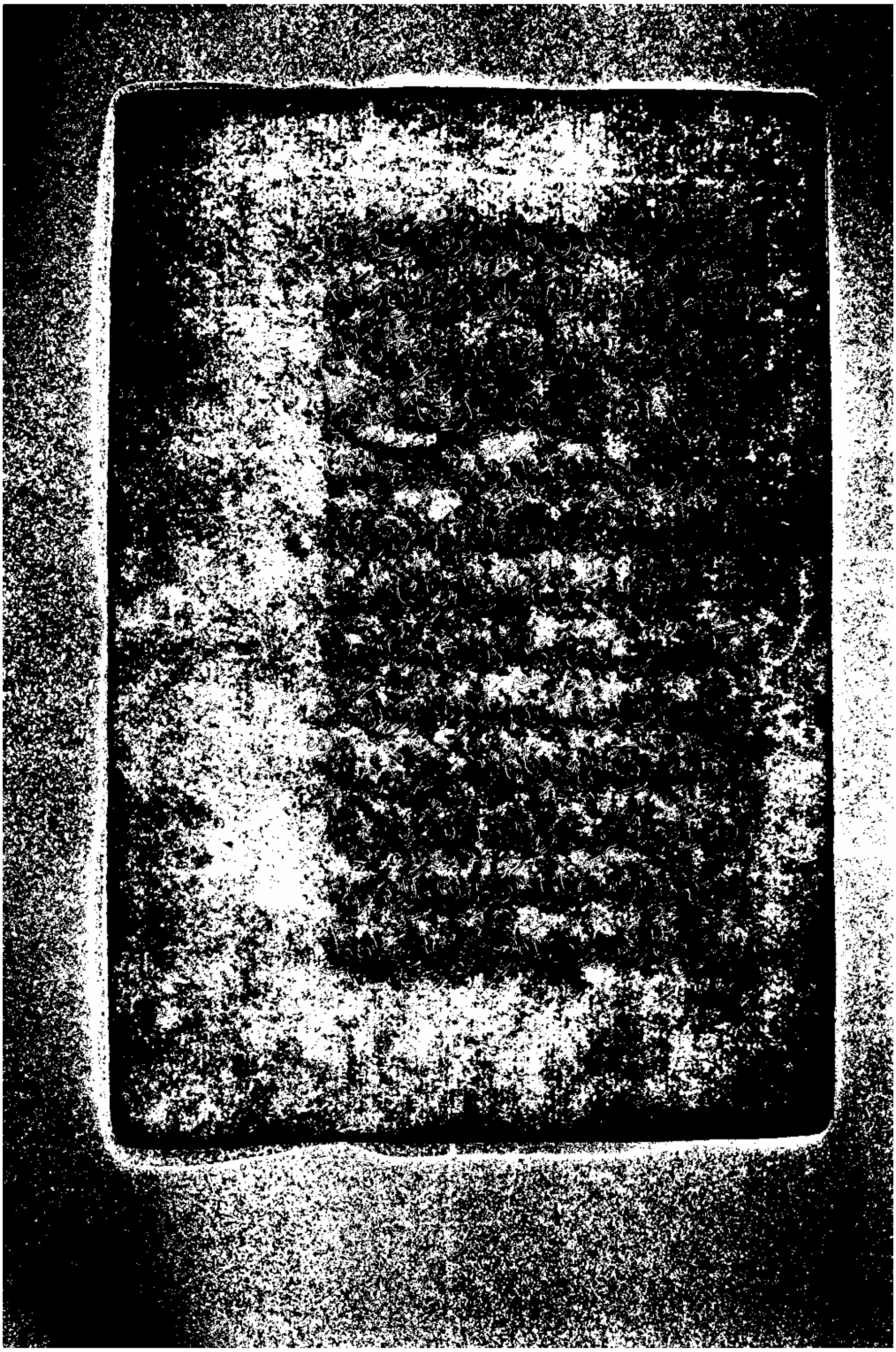


Figure B.23: A complete output of recto removal stage (differenced reflected recto) after contrast adjustment.

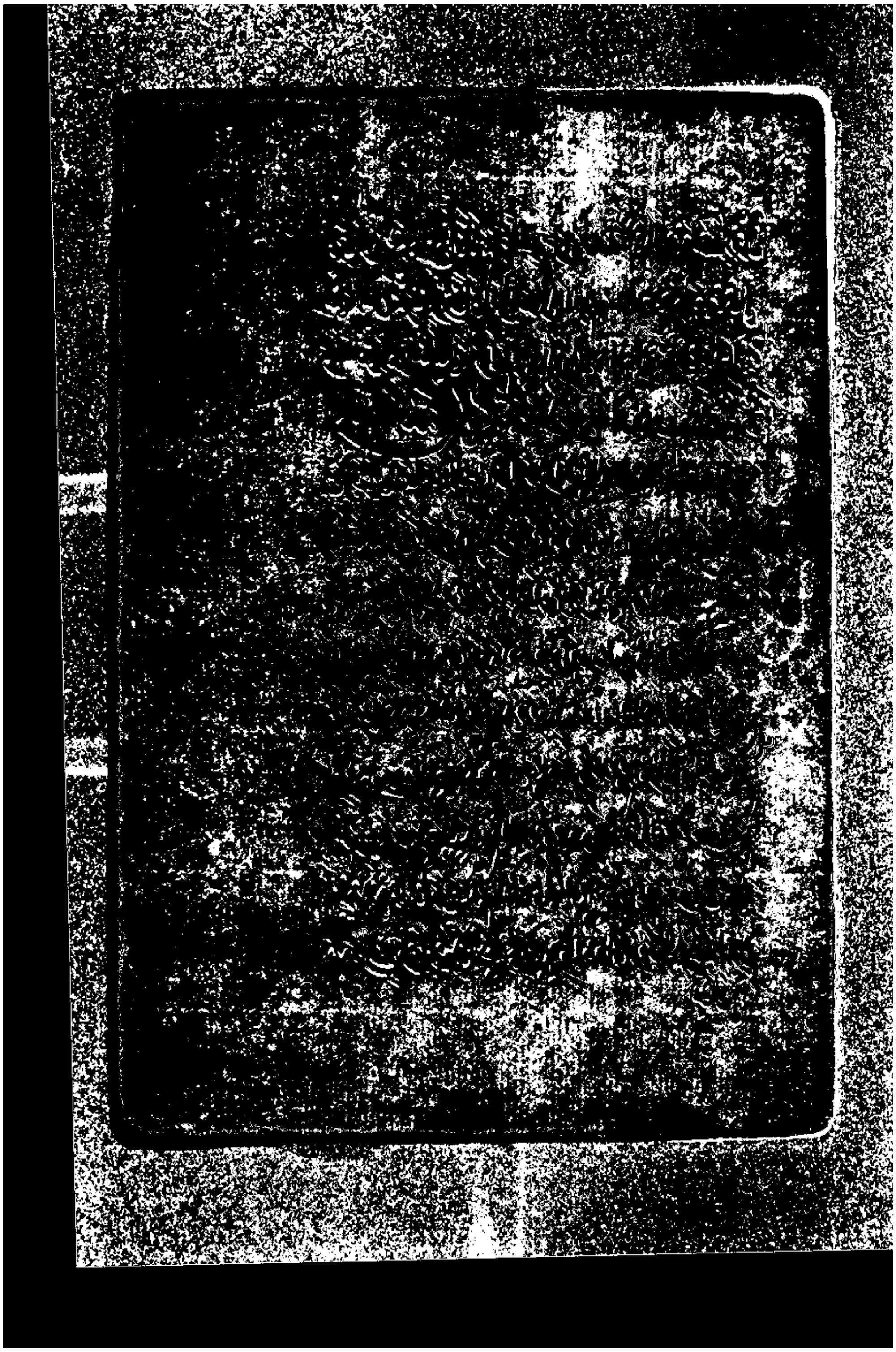


Figure B.24: A complete output of recto removal stage (differenced registered reflected verso) after contrast adjustment.

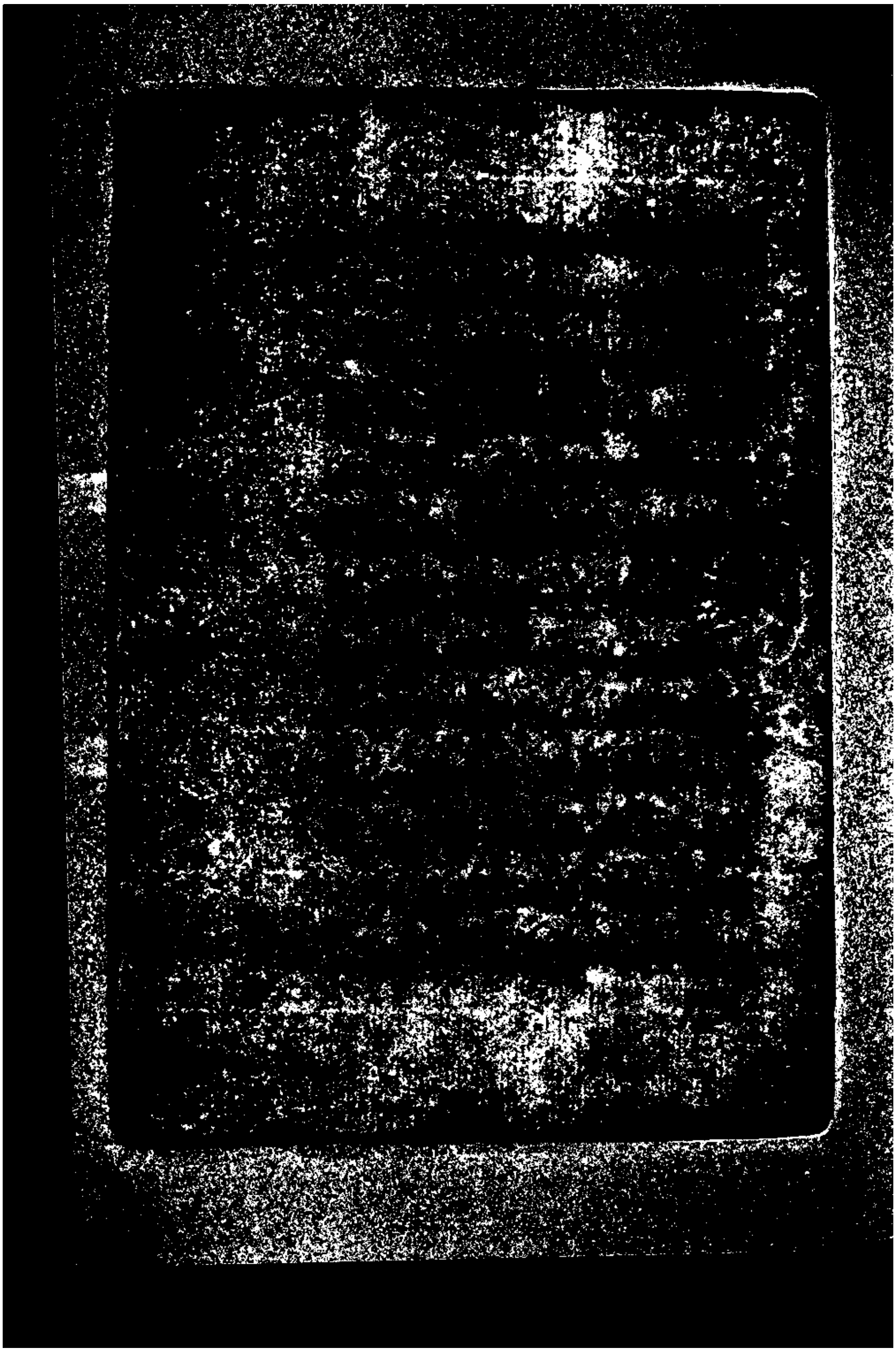


Figure B.25: The complete result of performing AND operation between B.23 and B.24 images.

Appendix C

Application Program Demo

We have included the MATLAB Compiler Runtime (MCR) with our deployed application; it is a standalone set of shared libraries that enable the execution of MATLAB files, even on computers without an installed version of MATLAB).

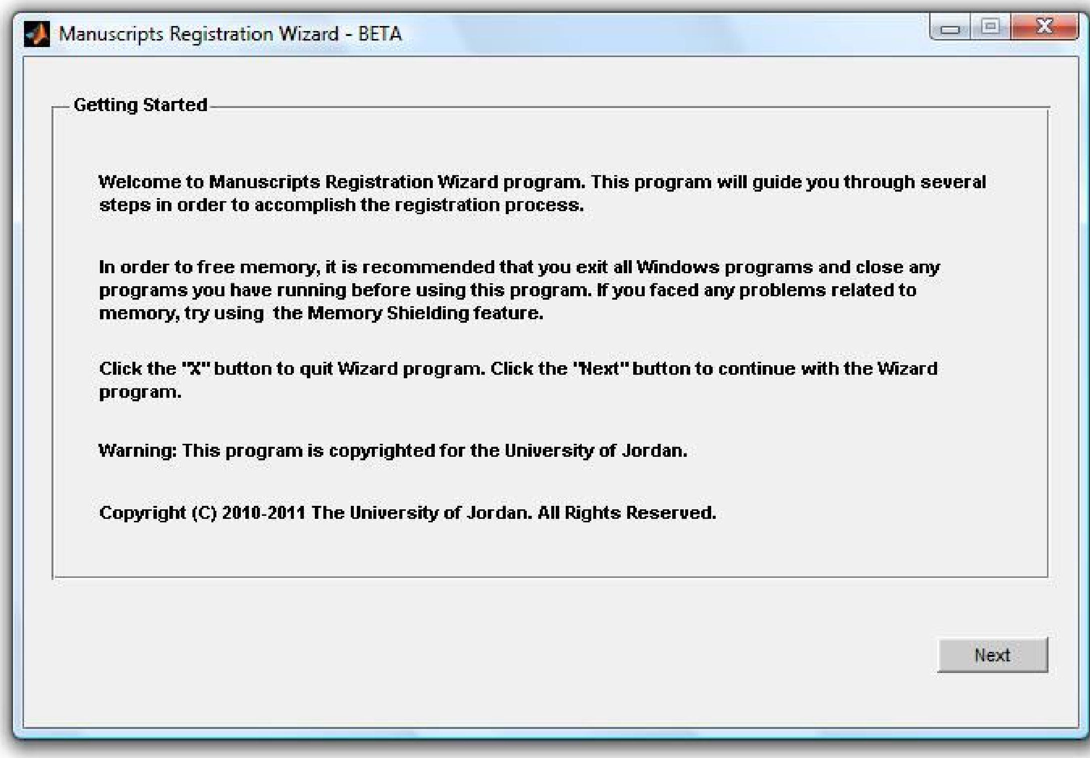


Figure C.1: The opening window

The main interface has a window for rendering of the input images and a set of controls of checkboxes and bush buttons. The application is operated a step at a time to trace all the main processing stages sequentially.

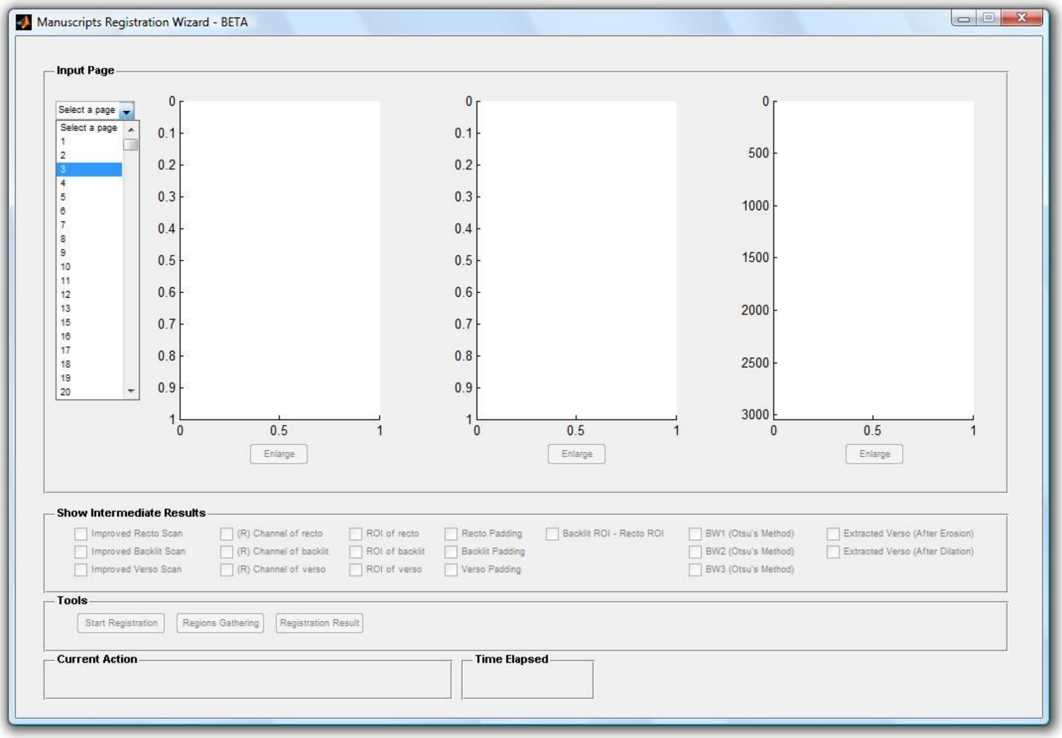


Figure C.2: The main GUI window.

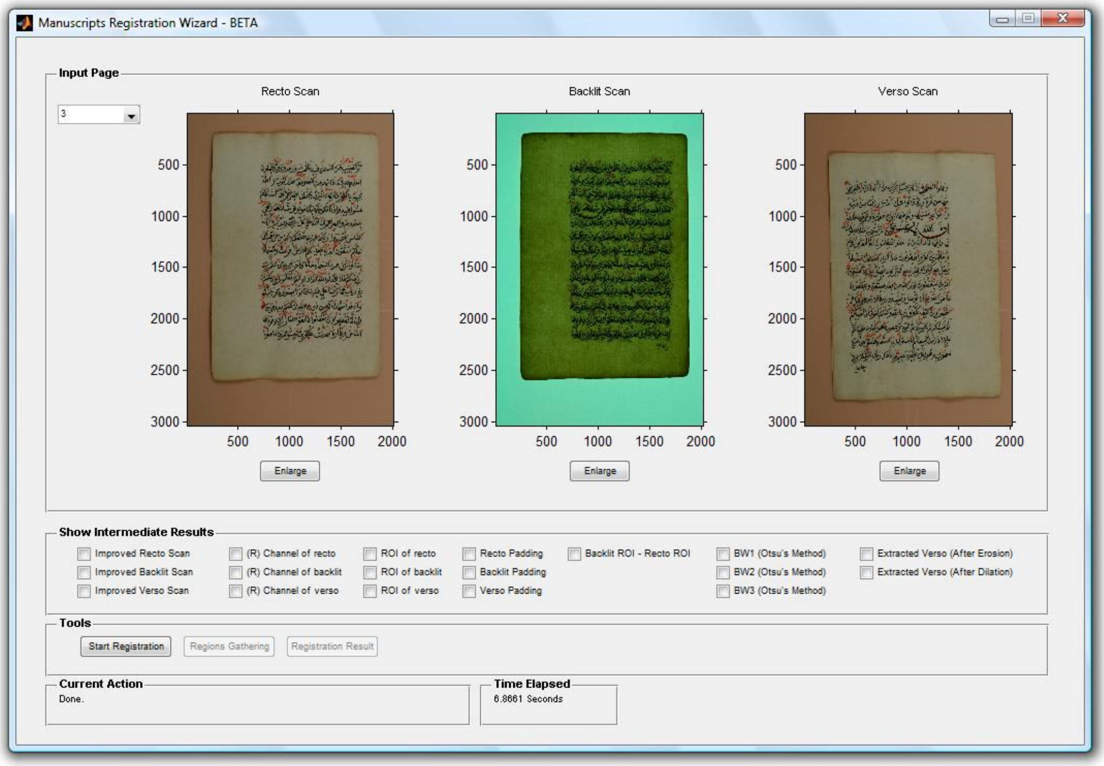


Figure C.3: The main GUI window after selecting a page from the 'Mahdiyya' copy of the Qur'an, with several options enabled to be selected.

Figure C.4: Message box appears when commencing the process of transmitted recto image cropping.

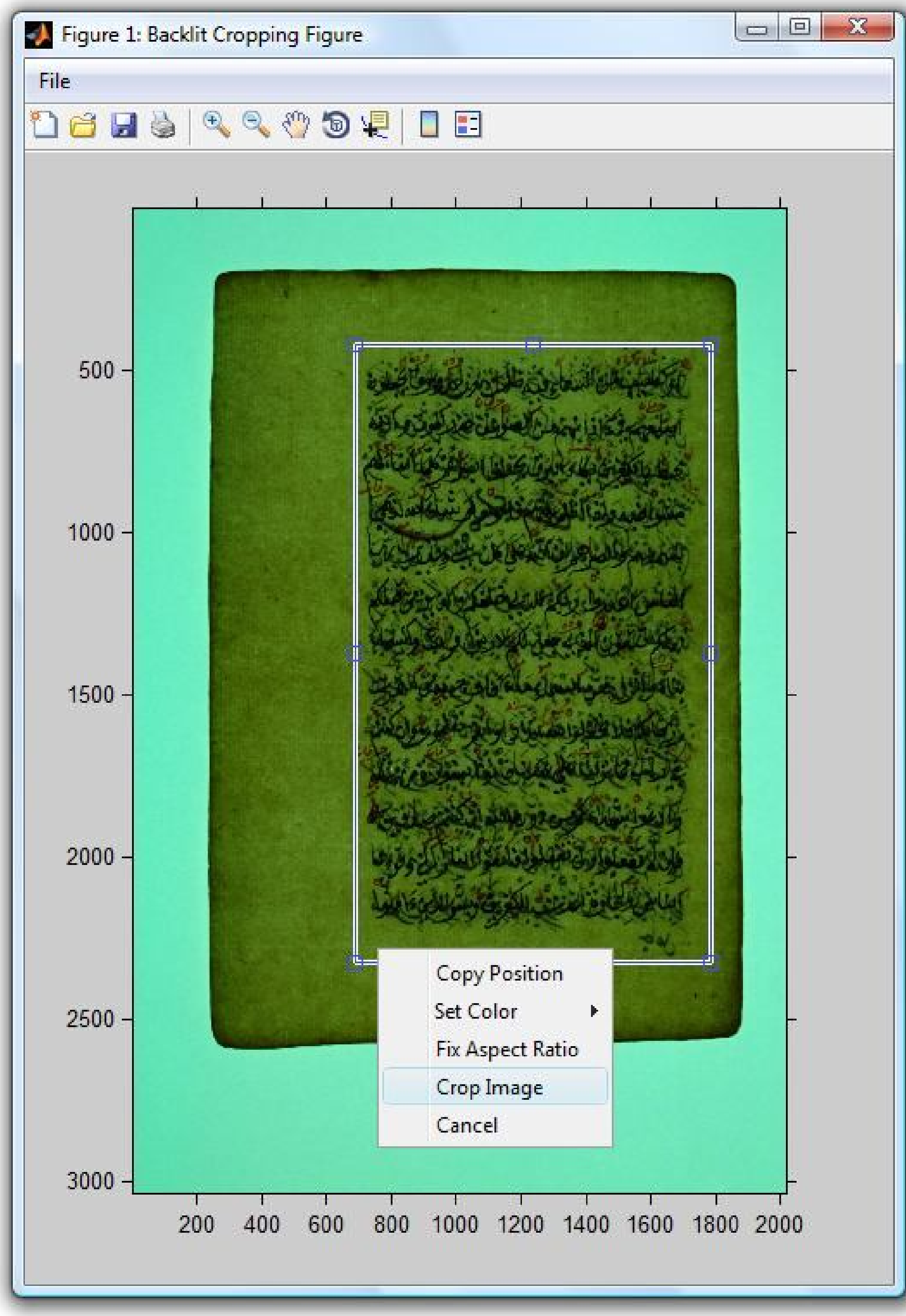


Figure C.5: Cropping window of the transmitted recto image.

Figure C.6: Message box appears when commencing the process of reflected verso cropping.

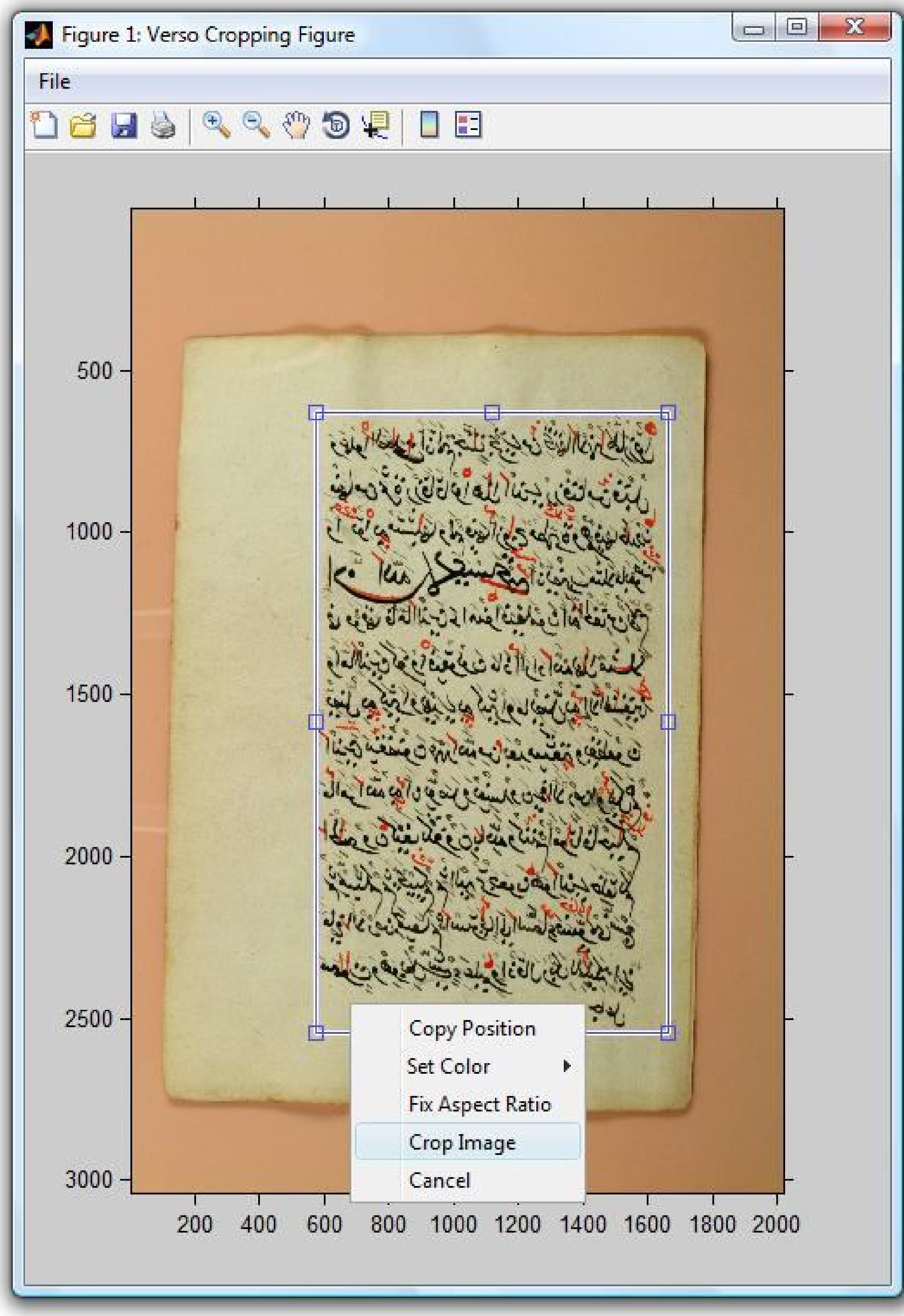


Figure C.7: Cropping window of the reflected verso image.

Figure C.8: Warning message that asks you to wait.

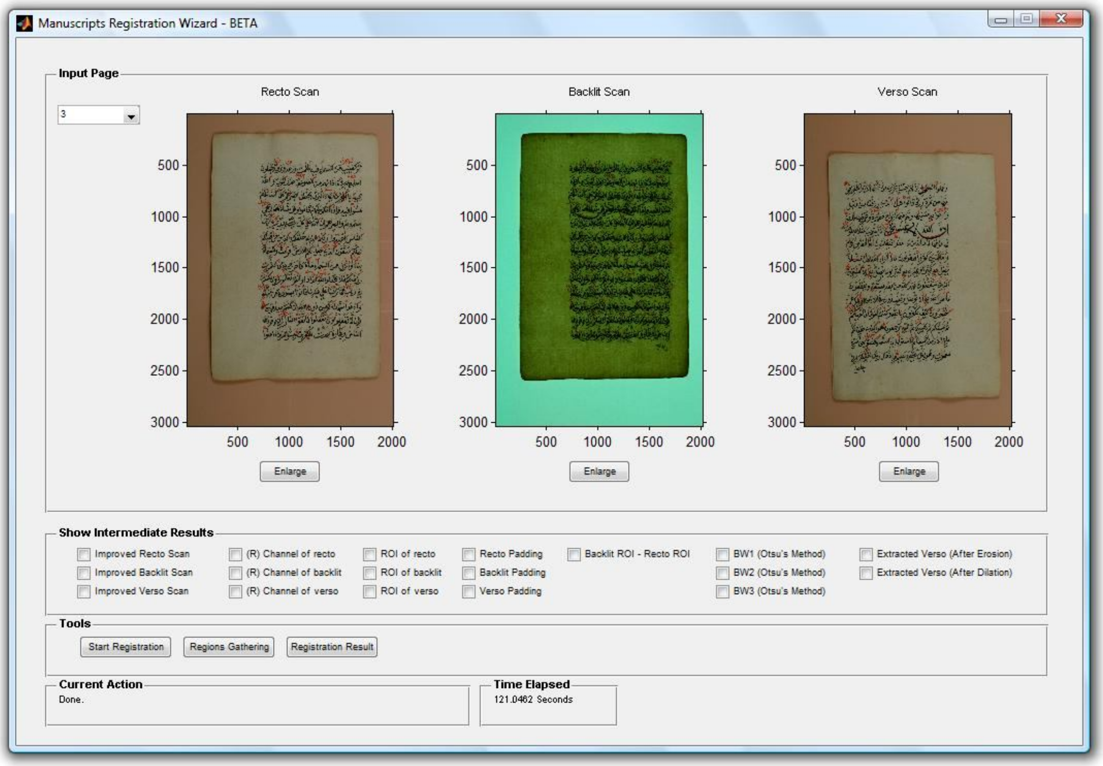


Figure C.9: Main Window after finishing the process of TR and RV registration.

If you encountered problems ensuring that you don't have the memory needed to run our deployed application, use MATLAB Memory Shielding to ensure that you obtain the maximum amount of contiguous memory to run the application successfully. This problem often occurs because the tested data set is large and the computer available to you has limited resources. The process reserves memory to ensure resource-intensive applications are allocated proper run-time resources. The memory shielding feature is installed with the MATLAB Compiler Runtime (MCR).

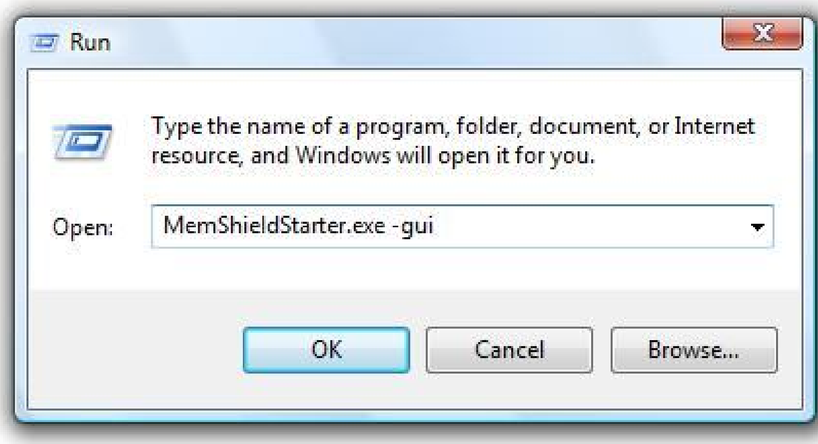


Figure C.10: Running memory shield.

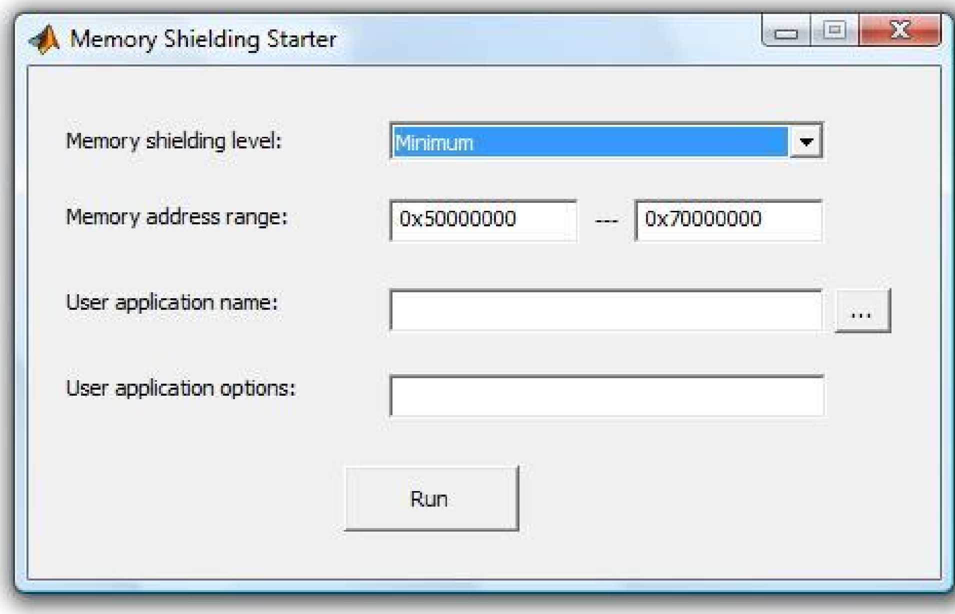


Figure C.11: Memory Shielding window.

استخراج العلامات المائية من الوثائق باستخدام الإضاءة الخلفية و معالجة الصور الرقمية

إعداد
جمال نصّار عمر سعيد

المشرف
الدكتور حازم الحياري

ملخص

تقدم هذه الرسالة نهجا جديدا يجمع بين مختلف تقنيات معالجة الصور الرقمية، محددة ومخصصة لاستخراج العلامات المائية المضمنة في الوثائق القديمة المكتوبة بخط اليد على الوجهين، من خلال القضاء على التشوهات التي تعترض العلامات المائية كالكتابات الموجودة على وجه و ظهر الورق، حيث تشكل تشويشا يؤثر بدوره على وضوحها.

لقد تم التقاط صور الوثائق رقميا بواسطة تقنية الإضاءة الخلفية، حيث تم التقاط ثلاث صور لكل ورقة من أوراق الوثائق، من الأمام والخلف و صورة عند تعريضها إلى الإضاءة الخلفية.

تم دمج المطابقة التلقائية المبنية على أساس الكثافة اللونية من أجل محاذاة الصور الأمامية والخلفية من الوثائق القديمة حيث تهدف هذه العملية إلى إزالة الكتابات الخلفية من الصور التي تم تعريضها إلى المسح الضوئي من أجل تقوية إشارة العلامات المائية. إن مرحلة مطابقة الصورة تتألف من نموذجين :المطابقة في النطاق المكاني و تليها المطابقة في القطب اللوغرثمي. من بين خوارزميات المطابقة الممكنة و المتاحة، ركزنا على تقنية الارتباط حيث اعتمدت هذه المرحلة على خوارزمية تقدير عوامل الارتباط.

لقد قمنا بإدراج تقنية المطابقة ضمن الإجراء المقترح بعد إجراء بعض الاختبارات التي قيمت أداء العملية بين الصور التي تم تعريضها للإضاءة الخلفية و الصورة الخلفية العادية من أجل استخراج العلامات المائية من أوراق الوثائق القديمة بنتائج مناسبة أو مقبولة بالرغم من وجود مشكلات ذات صلة بالتشوهات الناجمة عن طي الورق أو التكويم.

و قد أظهرت النتائج أن تقنية المطابقة قد حققت متانة جيدة لاحتواء التشوهات الهندسية التي تشمل الإزاحة والميلان والتحجيم. حيث أخذ بعين الاعتبار قياس جودة نتائج عملية المطابقة بالتقييم الكمي. من الناحية الأخرى، تم أيضا تقييم العلامات المائية المستخرجة نوعيا بواسطة العين البشرية المجردة.